# NON-STANDARD ANALYSIS

LECTURES

ON

A. ROBINSON'S THEORY OF INFINITESIMALS

AND

INFINITELY LARGE NUMBERS

BY

W. A. J. LUXEMBURG

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA

# NON-STANDARD ANALYSIS

Lectures

on

A. Robinson's Theory of Infinitesimals and Infinitely Large Numbers

by

W. A. J. Luxemburg



Second corrected edition, 1964

Mathematics Department
California Institute of Technology
Pasadena, California. 1966.

# PREFACE

The present lecture notes have grown from a series of three lectures which were given by the author at the California Institute of Technology in December 1961. The purpose of these lectures was to give a discussion of A. Robinson's theory of infinitesimals and infinitely large numbers which had just appeared in print under the title "Non-Standard Analysis". The title "Non-Standard Analysis" refers to the fact that this theory is an interpretation of analysis in a non-standard model of the arithmetic of the real numbers.

The notes contain six chapters. In Chapter 1 a detailed discussion is given of the construction of a non-standard model of the real number system in the form of an ultrapower, a concept which was introduced by Frayne, Scott and Tarski. This non-standard model is chosen in such a way, in order to enable us to compare Robinson's theory with the theory of infinitesimals and infinitely large numbers which was recently given by Schmieden and Laugwitz.

Chapters 2 and 3 are mainly concerned with the non-standard interpretation of the main principles of analysis. In addition, we develop a large part of the elementary theory of real functions using infinitesimals and infinitely large numbers. An important role in this theory plays the fundamental fact that every finite number, i.e., not infinitely large number, is infinitely close to a unique real number.

In these chapters the reader may find some new proofs of some well-known theorems. For instance, the author believes that the proof of the Bolzano-Weierstrass theorem (Theorem 1.7 of Chapter 3), Bolzano's intermediate value theorem for continuous functions (Theorem 5.1 of Chapter 3) and the proof of the mean-value theorem of the differential calculus (Theorem 10.2 of Chapter 3) are new.

In Chapter 4, some parts of the theory functions of several variables are treated in non-standard analysis.

Except for a few remarks about the theory of Riemann integration, no attempt has been made to discuss the theory of integration from a non-standard point of view. We hope to return to this question in the future.

Chapter 5 is concerned with the elementary theory of distributions which was recently given by Mikusinski and Sikorski. It is shown that a distribution may be represented as a generalized point function provided the operations on those functions are defined relative to a certain equivalence relation. Some possible realizations of the Dirac-delta distribution in the form of such generalized point functions are discussed. Needless to say that the given treatment of the theory of distributions is far from being complete.

The last chapter is devoted to a presentation of the properties of the general theory of ultrapowers of the real number system. In order to show how such general models can be used in analysis a number of applications are given. For instance, the Hahn-Banach extension theorem can be proved by means of an ultrapower construction of the reals. In place of Zorn's lemma, this proof bases the validity of this important extension theorem on the apparently weaker hypothesis that every filter is contained in an ultrafilter or, what is the same, the prime ideal theory for Boolean algebras.

For the sake of convenience, at certain places in the text we have made use of the following symbols: $\Rightarrow$; $\Leftrightarrow$; $(\exists \cdot)$; and $(\forall \cdot)$, which denote respectively the logical connectives "if..., then"; "if and only if..."; "there exists..."; and "for all...".

Every chapter is divided up in sections. In every section we have renumbered the theorems; the number of the section precedes the number of the theorem. If we refer to a theorem in a chapter to which it belongs, then the chapter is not quoted. In the other case it is quoted.

It is a great pleasure for me to thank my colleague Professor C. R. DePrima for the many stimulating discussions I had with him during the preparation of these notes.

Part of this work was supported by National Science Foundation Grant G-19914.

W. A. J. Luxemburg

Pasadena, January 1962.

The second edition differs little from the first edition. Some errors are corrected, section 4 of Chapter II was newly written and section 6 was added to Chapter 6.

For the metamathematical background of Robinson's theory we refer the reader to: Abraham Robinson, Introduction to Model Theory and the Metamathematics of Algebra, Studies in Logic and the Foundations of Mathematics, Amsterdam 1963.

W. A. J. Luxemburg

Pasadena, March 1964.

TABLE OF CONTENTS

PREFACE ii

CHAPTER 1. A NON-STANDARD MODEL FOR ANALYSIS

1. Some Historical Remarks; References 1
2. Filters and Ultrafilters 3
3. Some Remarks about Rings and Fields 6
4. Construction of a Non-Standard Model for Analysis 15
5. Infinitesimals, Finite and Infinitely Large Numbers 20
6. The Interval Topology of $R^*$ 23
7. Non-Standard Extensions 25

CHAPTER 2. THE THEORY OF LIMITS OF SEQUENCES

1. Convergent Sequences 38
2. The Algebra of Limits 41
3. Cauchy's Criterion for Convergence 42
4. The Existence of Generalized Limits in the Sense of S. Banach 44
5. A Theorem of G. Pólya and G. Szegö 46
6. The Theory of Limits of Double Sequences 49
7. The Theory of Infinite Series 55

CHAPTER 3. THE THEORY OF REAL FUNCTIONS

1. The Topology of $R$ in Non-Standard Analysis 63
2. Limits of Functions 68
3. Continuity 73
4. Properties of Real Continuous Functions 75
5. The Intermediate Value Theorem for Continuous Functions 77
6. Uniform Continuity 79
7. Additive Functions 80
8. Differentiation of Functions of a Real Variable 83
9. The Chain Rule 85
10. Mean-Value Theorems 86
11. Uniform Convergence; Equicontinuity 90
12. The Riemann Integral 100

CHAPTER 4. FUNCTIONS OF SEVERAL VARIABLES

1. A Non-Standard Model for $R_p$ 102
2. Non-Standard Extensions 104
3. Convergence of Sequences in $R_p$ 106
4. The Topology of $R_p$ 107
5. Limits of Functions 108
6. Continuity 108
7. Uniform Continuity 109
8. Differentiation of Functions of Several Variables 110
9. The Mean-Value Theorem 112

CHAPTER 5. THE THEORY OF DISTRIBUTIONS


1. Quasi-Standard Functions 115
2. Dirac-Delta Functions 121
3. Definition of Distributions of Finite Order 124
4. Algebraic Operations on Distributions 129
5. Derivation of Distributions 130


CHAPTER 6. ULTRAPOWERS IN ANALYSIS


1. Introduction 134
2. General Non-Standard Models of R in the Form of Ultrapowers 134
3. The Hahn-Banach Extension Theorem 141
4. Tarski's Extension Theorem for Measures on Boolean Algebras 143
5. Nikodým's Theorem about the Existence of Strictly-Positive Non-Archimedean Measures on Boolean Algebras 147
6. The Heine-Borel Covering Theorem 150

CHAPTER I.

# A NON-STANDARD MODEL FOR ANALYSIS

## 1. Some Historical Remarks; References.

The use of infinitesimals in analysis was strongly advocated by Leibniz and readily accepted by Euler. Infinitesimals were abolished after the advent of Cauchy's methods which made analysis a rigorous branch of mathematics.

Except for a few instances such as du Bois-Reymond's calculus of infinities and Hahn's work on non-archimedean fields no complete rigorous theory of infinitesimals and infinitely large quantities was put forward until recently. In fact, two such theories were offered, one by Schmieden and Laugwitz in 1958 and another one by Robinson in 1961.

These two theories are very much different from each other. Schmieden and Laugwitz arrived at their theory by means of a new approach to Cantor's definition of the system of real numbers. The generalized system of numbers they obtained is an ordered (= partially ordered) ring with divisors of zero and which contains infinitesimals and infinitely large numbers.

Robinson's theory is based on the metamathematical fact that the system of real numbers $R$ is incomplete. Thus, there exist proper extensions of $R$ which possess all properties of $R$ that are formulated in the lower predicate calculus in terms of some given set of number theoretic relations such as addition, multiplication and equality. A fact which was observed earlier by Skolem for the system of natural numbers $N(n \geq 0)$.

Proper extensions of non-complete theories are often referred to as (strong) non-standard models. A non-standard model for the system of real numbers has the feature of being a non-archimedean totally ordered field which contains a copy of the real number system. Non-standard models can be constructed as ultrapowers, a metamathematical concept which was recently introduced by T. Frayne, D. Scott and A. Tarski.

We shall construct a non-standard model of analysis in the form of an ultrapower. Furthermore, the ultrapower is chosen in such a way that it will throw some light on the relation between the two theories.

Finally, the reader should take notice of the fact that there are an infinite number of non-isomorphic non-standard models for the theory of real numbers. One may speculate that by choosing a particular model one could analyze certain parts of analysis more closely. We shall illustrate this method in the final chapter. For the general theory presented in all but the last chapter any non-standard model for $R$ will do as well as another.

In order to make this report as self-contained as possible we have included the basic definitions and properties of the theory of filters and proofs of some theorems about ordered fields.

We conclude this section with the following list of articles. In the articles [1], [6] and [7] the reader may find some additional information about the history of the subject.

1. Schmieden, C. and D. Laugwitz, Eine Erweiterung der Infinitesimalrechnung, Math. Zeitschr., 69, 1-39 (1958).

2. Laugwitz, D., Eine Einführung der δ-Funktionen, Sitzungsber. Bayerische Akad. der Wissenschaften, 4, 41-59 (1959).

3. Laugwitz, D., Anwendungen unendlich kleiner Zahlen I. Zur Theorie der Distributionen, J. reine angew. Math., 207, 53-60 (1961).

4. Laugwitz, D., Anwendungen unendlich kleiner Zahlen II. Ein Zugang zur Operatorenrechnung von Mikusinski, J. reine angew. Math., 208, 22-34 (1961).

5. Bopp, F., Lorentzinvariante Wellengleichungen für Mehrbahnsysteme, Sitzungsber. Bayerische Akad. der Wiss. 167-225 (1958).

6. Erdélyi, A., An extension of the concept of real number, To be published in the Proc. Fifth Canad. Math. Congress, Montreal (1961).

7. Robinson, A., Non-standard analysis, Proc. Nederl. Akad. Wetensch. 64 (1961) (= Ind. Math., 23, 432-440 (1961)).

8. Frayne, T., D. Scott and A. Tarski, Reduced Products, Amer. Math. Soc. Notices, 5, 673-674 (1958).

9. Erdös, P., L. Gillman and M. Henriksen, An isomorphism theorem for real-closed fields, Annals of Mathematics 61, 542-544 (1955).

## 2. Filters and Ultrafilters

DEFINITION 2.1 (Filter). *A non-empty set* $\mathfrak{F}$ *of subsets of a non-empty set* $X$ *is called a filter if it has the following filter properties*:

(F1) *If* $E \in \mathfrak{F}$ *and* $F \supseteq E$, *then* $F \in \mathfrak{F}$.

(F2) *If* $E, F \in \mathfrak{F}$, *then* $E \cap F \in \mathfrak{F}$.

(F3) *The empty subset of* $X$ *is not an element of* $\mathfrak{F}$.

From property (F3) it follows that a filter cannot degenerate in the set of all subsets of $X$. Furthermore, (F2) and (F3) imply that every finite set of elements of $\mathfrak{F}$ has a non-empty intersection. Since $\mathfrak{F} \neq \phi$, (F1) implies $X \in \mathfrak{F}$.

EXAMPLES 1. If $\mathcal{F} = \{X\}$, then $\mathcal{F}$ is a filter.

2. If $Y$ is a non empty subset of $X$, then $\{E : Y \subseteq E\}$ is a filter. This filter is called the principal filter on $X$ generated by $Y \subseteq X$.

3. (Fréchet filter). The set of all non-empty subsets of $X$ whose complements are finite is a filter on $X$. This filter is called the Fréchet filter of $X$ and is denoted by $\mathcal{F}r$.

DEFINITION 2.2 (Free filter). A filter $\mathcal{F}$ is called free if $\bigcap\{E : E \in \mathcal{F}\} = \phi$.

EXAMPLES 1. A principal filter is not free.
2. If $X$ is infinite, then the Fréchet filter of $X$ is a free filter. Indeed, for every $x \in X$ we have $X - \{x\} \in \mathcal{F}r$. The reader should also observe that the following converse holds: If a set $X$ has a free filter, then $X$ is infinite.

DEFINITION 2.3 Let $\mathcal{F}$ and $\mathcal{F}'$ be two filters on $X$. Then $\mathcal{F}'$ is called finer than $\mathcal{F}$ if $\mathcal{F} \subseteq \mathcal{F}'$, i.e. $(\forall E)(E \in \mathcal{F} \Rightarrow E \in \mathcal{F}')$. If, moreover, $\mathcal{F} \neq \mathcal{F}'$, then $\mathcal{F}'$ is called strictly finer than $\mathcal{F}$.

The set of all filters on $X$ is ordered by the relation "$\mathcal{F}'$ is finer than $\mathcal{F}$." Indeed, this relation is the induced relation of the inclusion relation in the power set of the power set of $X$. In the following for "$\mathcal{F}'$ is finer than $\mathcal{F}$" we shall write "$\mathcal{F} \leq \mathcal{F}'$."

DEFINITION 2.4 (Ultrafilter). A filter $\mathcal{F}$ on $X$ is called an ultrafilter if it is not properly contained in any other filter, or equivalently, if there does not exist a filter on $X$ which is strictly finer than $\mathcal{F}$.

Ultrafilters are maximal elements of the family of all filters ordered by the relation "$\mathcal{F} \leq \mathcal{F}'$."

Since it is easy to see that the set of all filters on X ordered by $\mathcal{F} \leq \mathcal{F}'$ has the property that every chain has an upper bound the following theorem follows immediately from Zorn's lemma.

THEOREM 2.1. *For every filter* $\mathcal{F}$ *on* X *there exists an ultrafilter* $\mathcal{U}$ *on* X *such that* $\mathcal{F} \leq \mathcal{U}$.

THEOREM 2.2. *Let* $\mathcal{U}$ *be an ultrafilter on* X. *If* E *and* F *are two subsets of* X *such that* $E \cup F \in \mathcal{U}$, *then either* $E \in \mathcal{U}$ *or* $F \in \mathcal{U}$.

PROOF. We shall give an indirect proof. Assume that $E \notin \mathcal{U}$ and $F \notin \mathcal{U}$ and $E \cup F \in \mathcal{U}$. Let $\mathcal{F}$ be the set of all subsets Y of X such that $E \cup Y \in \mathcal{U}$. It is easy to verify that $\mathcal{F}$ is a filter. Since $F \in \mathcal{F}$, we have that $\mathcal{F}$ is strictly finer than $\mathcal{U}$. This contradicts the fact that $\mathcal{U}$ is an ultrafilter and the proof is completed.

THEOREM 2.3. *Let* $\mathcal{U}$ *be an ultrafilter on* X. *If* $\{E_i : 1 \leq i \leq n\}$ *is a finite family of subsets of* X *such that* $\bigcup\{E_i : 1 \leq i \leq n\} \in \mathcal{U}$, *then there exists at least one index* i $(1 \leq i \leq n)$ *such that* $E_i \in \mathcal{U}$. *In particular, if* $\bigcup\{E_i : 1 \leq i \leq n\} = X$, *then the conclusion of the theorem holds*.

PROOF. This theorem follows immediately from the preceding theorem by induction on n.

The following theorem gives an important characteristic property of an ultrafilter.

THEOREM 2.4. *A filter* $\mathcal{F}$ *on* X *is an ultrafilter if, and only if, for every subset* Y *of* X *either* $Y \in \mathcal{F}$ *or* $X - Y \in \mathcal{F}$.

PROOF. If $\mathcal{F}$ is an ultrafilter, then the result follows immediately from Theorem 2.2. Conversely, assume that $\mathcal{F}$ is a filter such that for all subsets $Y$ of $X$ either $Y \in \mathcal{F}$ or $X - Y \in \mathcal{F}$. If $\mathcal{F}$ is not an ultrafilter, then there exists a subset $Y$ of $X$ such that (i) $Y \notin \mathcal{F}$ and (ii) $E \cap Y \neq \phi$ for all $E \in \mathcal{F}$. Hence, $Y \notin \mathcal{F}$ implies that $X - Y \in \mathcal{F}$. This contradicts (ii).

THEOREM 2.5. _If_ $X$ _is infinite, then there exists a free ultrafilter on_ $X$.

PROOF. If $X$ is infinite, then the Fréchet filter of $X$ is free. From Theorem 2.1 it follows that there exists an ultrafilter $\mathcal{U}$ such that $\mathcal{F}r \leq \mathcal{U}$. Since $\bigcap \mathcal{U} \subseteq \bigcap \mathcal{F}r$ we obtain that $\mathcal{U}$ is a free ultrafilter.

REMARK. The non-free or _fixed_ ultrafilters on a set $X$ can be easily characterized as the principal filters generated by the subsets $\{x\}$ of $X$, where $x \in X$. It may be of interest to remark that since the fixed ultrafilters of a set $X$ are those filters whose elements contain a fixed element of $X$, the existence of such ultrafilters follows readily. The existence of free ultrafilters on infinite sets, however, has never been shown then by using some form of the axiom of choice. Incidentally, the converse of Theorem 2.5 holds also.

## 3. Some Remarks about Rings and Fields.

For terminology not explained in this section the reader should consult N. H. McCoy, Rings and Ideals, Carus Mathematical Monograph 8 and N. Bourbaki, Livre II Algèbre, Chap. VI, Groupes et corps ordonnés, Hermann No. 1179.

In what follows, $A$ will always denote a *commutative* ring which has an identity, i.e., an element 1, necessarily *unique*, such that $1 \cdot x = x$ for all $x \in A$.

An element $x \in A$ is called *regular* if it has a multiplicative inverse, i.e., a necessarily *unique* element, denoted by $a^{-1}$ or $1/a$, such that $a\,a^{-1} = 1$.

DEFINITION 3.1 (*Ideal*). *A subring* $I$ *of* $A$ *is called an ideal if* $I \neq A$ *and* $a \in A$ *implies* $a\,x \in I$ *for all* $x \in I$.

If $I$ is an ideal, then $1 \notin I$, since $I \neq A$. More generally, *every regular element does not belong to any ideal*.

DEFINITION 3.2 (*Homomorphism*). *Let* $A$ *and* $A'$ *be two commutative rings with identity. A mapping* $h$ *of* $A$ *into* $A'$ *is called a ring homomorphism if (i)* $h(x + y) = h(x) + h(y)$ *for all* $x, y \in A$ *and* (ii) $h(xy) = h(x)\,h(y)$ *for all* $x, y \in A$.

If $h$ is a homomorphism of $A$ into $A'$, then $h(0)$ is the zero-element of $A'$. If, moreover, $h$ is an onjection of $A$ onto $A'$, i.e., $h(A) = A'$, then $h(1)$ is the identity of $A'$. Thus a *homomorphic image of a ring with identity is a ring with identity*.

A homomorphism $h$ of $A$ onto $A'$ such that $x \neq y$ implies $h(x) \neq h(y)$ is called an *isomorphism* and $A$ and $A'$ are called *isomorphic*.

The *kernel* $I$ of a non-zero homomorphism of $A$ into $A'$, i.e., the set of all $x \in A$ such that $h(x) = 0$, is an ideal. In fact, $I$ is the kernel of the *canonical homomorphism* of $A$ onto the residue class ring $A/I$. Conversely, every ideal $I \subset A$ is the kernel of the canonical homomorphism of $A$ onto $A/I$. If $h$ is a homomorphism of $A$ onto $A'$, then $A/I$ and $A'$ are isomorphic.

DEFINITION 3.3 (<u>Maximal ideal</u>). <u>An ideal</u> $I$ <u>of</u> $A$ <u>is called maximal if there does not exist an ideal</u> $I' \subset A$ <u>such that</u> $I \subseteq I'$ <u>and</u> $I \neq I'$.

The property of an ideal to be maximal can be expressed algebraically as follows:

THEOREM 3.1. <u>An ideal</u> $I \subset A$ <u>is maximal if and only if</u> $A/I$ <u>is a field</u>.

PROOF. Assume that $A/I$ is a field. Then, $a \notin I$ and $b \in A$ implies there exists an element $c \in A$ such that $b - a\,c \in I$. Hence, any ideal $I'$ containing $I$ as a subset and $a$ as an element is equal to $A$. We conclude that $I$ is maximal.

Conversely, assume that $I$ is maximal and that $a \notin I$. Then there exists an element $b \in A$ such that $ab - 1 \in I$. Hence $a^2b - a \in I$. From $a \notin I$ it follows that $a^2 \notin I$. Thus any ideal $I'$ containing $a$ and $I$ contains $I$ properly. Hence, $I' = A$, since $I$ is maximal. We conclude that given $b \in A$, there exists an element $c \in A$ such that $b - ac \in I$, i.e., $A/I$ is a field.

Consider the set of all ideals of $A$ ordered by inclusion. The union of any non-empty chain of ideals is an ideal. Hence, Zorn's lemma applies.

THEOREM 3.2. <u>Every ideal</u> $I \subset A$ <u>is contained in some maximal ideal</u>. In particular, every non-regular element of $A$ belongs to some maximal ideal.

THEOREM 3.3. *If* $A$ *has the property that the set of all non-regular elements of* $A$ *forms an ideal* $I$, *then* $I$ *is maximal*.

PROOF. If $I$ is not maximal, then there exists an element $a \in A$ and an ideal $I'$ such that $a \notin I$, $I \subseteq I'$ and $a \in I'$. But $a \notin I$ implies that $a$ is regular. Hence $a \notin I'$ and a contradiction is obtained.

We proceed with the introduction of the important concept of an ordered ring.

DEFINITION 3.4 (*Ordered ring*). *A ring* $A$ *is called an ordered ring if* $A$ *is ordered and the following two conditions hold:*

(i) $x \leq y$ *implies* $x + z \leq y + z$ *for all* $z \in A$,

(ii) *for all* $x, y \in A$, $x \geq 0$ *and* $y \geq 0$ *implies* $xy \geq 0$.

If $A$ is a totally-ordered ring and $a \in A$, then $|a|$ denotes $\max(a, -a)$.

Assume now that $A$ is a totally-ordered *integral domain* (i.e., a totally-ordered ring *without divisors of zero*). Then,

(i) $1 > 0$ *and hence* $-1 < 0$. Indeed, $1 < 0$ implies $1.1 > 0$ which contradicts $1 < 0$.

(ii) $x \neq 0$ *implies* $x^2 > 0$. Indeed, since $A$ is an integral domain, $x \neq 0$ implies $x^2 \neq 0$. Since $A$ is totally-ordered we have either $x > 0$ or $x < 0$. In both cases, however, we may conclude that $x^2 > 0$. In particular, $-1$ *has no square root*.

(iii) *If* $x < y$, *then* $x^n < y^n$ *for all* $n \geq 1$. Thus a positive element has at most one positive root.

(iv) <u>If</u> $x > 0$, <u>then</u> $nx \neq 0$ <u>for all</u> $n \geq 1$. Indeed, A has no divisors of zero.

(v) A <u>contains a natural copy of the set of integers in the form of the elements</u> m.1, <u>where</u> m <u>is an integer</u>.

(vi) <u>If</u> F <u>is a totally-ordered field, then the elements</u> $m/n = (m.1)/(n.1)$, <u>where</u> m <u>is an integer and</u> n <u>a natural number</u> $\geq 1$ <u>constitutes a natural copy of the field of rationals</u> Q.

DEFINITION 3.5 (<u>Archimedean totally-ordered field</u>). <u>A totally-ordered field</u> F <u>is called archimedean if for every</u> $a \in F$ <u>there exists an element</u> $n \in N$ <u>such that</u> $|a| < n$.

The field of rationals and the field of real numbers are examples of totally ordered archimedean fields.

Before we conclude this section with a characterization of the archimedean property for totally-ordered fields we shall recall briefly Dedekind's definition of the system of real numbers.

Let Q again denote the field of rational numbers.

DEFINITION 3.6. (<u>Dedekind cut</u>). <u>A subset</u> a <u>of</u> Q <u>is called a Dedekind cut if (i)</u> $a \neq \emptyset$, <u>(ii)</u> $r \in a$ <u>and</u> $r' \in Q$ <u>and</u> $r' < r$ <u>imply</u> $r' \in a$. <u>(iii)</u> a <u>contains no largest element</u>.

If a is a Dedekind cut, then $r \notin a$, $r \in Q$ implies $r' < r$ for all $r' \in a$.

If $r \in Q$, then the set $a = \{r' : r' \in Q \text{ and } r' < r\}$ is a Dedekind cut and r is the least upper bound of A in Q. This Dedekind cut is called the <u>rational cut</u> defined by r. The Dedekind cut defined by $0 \in Q$ will be denoted by 0. Thus $0 = \{r : r \in Q \text{ and } r < 0\}$.

Two Dedekind cuts $a$ and $b$ are considered to be equal if they are equal as sets, i.e., if they contain the same rationals.

Let $R$ be the family of all Dedekind cuts of $Q$. Then the relation $a \leq b \Leftrightarrow a = b$ or $(\exists r)(r \in Q$ and $r \in b$ and $r \notin a)$ totally orders $R$. Thus for every $a \in R$ we have either $a < 0$ or $a = 0$ or $a > 0$.

Addition in $R$ can be introduced as follows:

If $a, b \in R$, then by $a + b$ we denote that subset of $Q$ which has the following property: $r \in a + b \Leftrightarrow (\exists r')(\exists r'')(r' \in a, r'' \in b$ and $r = r' + r'')$. It is easy to see that $a + b$ is again a Dedekind cut. Then it is easily shown that $R$ is a totally-ordered commutative group, with respect to this definition of addition , which contains the additive group of $Q$ by means of the rational cuts.

The introduction of multiplication is more complicated.

If $a, b \in R$ and $a \geq 0$, $b \geq 0$, then by $a.b$ or shortly $ab$ we denote that subset of $Q$ which has the following property: $r \in ab \Leftrightarrow (\exists r')(\exists r'')(r' \in a, r'' \in b$ and $r = r'r'')$. It is easy to see that $ab$ is a Dedekind cut. Multiplication for general elements of $R$ is then introduced as follows:

If $a, b \in R$, then $ab = -|a|\,|b|$ whenever $a < 0$ and $b \geq 0$, $ab = -|a|\,|b|$ whenever $a \geq 0$ and $b < 0$ and $ab = |a|\,|b|$ whenever $a < 0$ and $b < 0$.

Then $R$ turns out to be an archimedean totally-ordered field with the property that every non-empty subset of $R$ which is bounded above has a least upper bound. For further information about the construction of the real number system by means of Dedekind cuts we refer the reader to: W. Rudin, Principles of Mathematical Analysis, New York (1953).

Using Dedekind's definition of the real number system we shall prove now the following theorem which will be essential in the proof of Theorem 5.2.

THEOREM 3.4. <u>A totally-ordered field is archimedean if and only if it is isomorphic to a subfield of</u> R.

PROOF. Since every subfield of R is archimedean we have only to show that the condition is necessary. For this purpose assume that F is a totally-ordered archimedean field. If $x, y \in F$ and $x < y$, then there exists an element $n \in N$ such that $n > 1/(y-x)$. Let m be the smallest integer $> nx$. Then $x < m/n < y$. Hence, Q is dense in F. So that every element x of F is uniquely determined by the Dedekind cut $\{r : r \in Q \text{ and } r < x\}$. Thus F is embeddable in R in a unique way as a totally-ordered subset of R. Furthermore, if $x, y \in F$ and $q, r, s, t$ are rationals such that $q \leq x < r$, $s \leq y < t$, then $s + q \leq x + y < r + t$. Hence, addition in F like addition in R is uniquely determined by Dedekind cuts of Q. This holds also for multiplication as the reader can easily verify himself. This shows that the embedding of F into R is an isomorphism, which completes the proof of the theorem.

We shall conclude this section with an analysis of a problem the solution of which plays an important role in the proof of Theorem 5.2.

Let A be an ordered ring and let I be a proper ideal of A. We wish to know when the residue class ring A/I can be ordered in such a way that A/I is an ordered ring in the sense of Definition 3.4 and that the canonical mapping h of A onto A/I is order preserving.

The latter requirement suggests immediately to consider the following relation between the elements of $A/I$.

(*) $h(a) \leq h(b)$ <u>if and only if there exist elements</u> $x, y \in A$ <u>such that</u> $x \leq y$ <u>and</u> $x - a \in I$, $y - b \in I$.

Indeed, from (*) it follows immediately that if $a \leq b$, then $h(a) \leq h(b)$. Furthermore, $h(a) \leq h(b)$ if and only if $h(b - a) \geq 0$. If $h(a) \leq h(b)$, then by (*) there exist elements $x, y \in A$ such that $x \leq y$ and $x - a \in I$, $y - b \in I$, i.e., $y - x \geq 0$ and $(b - a) - (y - x) \in I$, or equivalently $h(b - a) \geq 0$. Conversely, if $h(b - a) \geq 0$, then by (*) there exists an element $z \in A$ such that $z \geq 0$ and $b - a - z \in I$. If $x \in A$ such that $a - x \in I$, then $b - (z + x) = b - (z + x - a + a) = (b - a) - z + a - x \in I$. Since $A$ is an ordered ring $z \geq 0$ implies $z + x \geq x$. Hence, $h(b) \geq h(a)$.

We shall prove now the following theorem.

THEOREM 3.5. <u>The relation</u> $h(a) \leq h(b)$ <u>between the elements of</u> $A/I$ <u>defined by</u> (*) <u>has the following properties</u>:

(a) <u>It is reflexive, i.e</u>, $h(a) \leq h(a)$ <u>for all</u> $a \in A$.

(b) <u>It is transitive, i.e.</u>, $h(a) \leq h(b)$ <u>and</u> $h(b) \leq h(c)$ <u>implies</u> $h(a) \leq h(c)$.

(c) <u>It satisfies conditions (i) and (ii) of Definition 3.4</u>.

(d) <u>It is antisymmetric, i.e.</u>, $h(a) \leq h(b)$ <u>and</u> $h(b) \leq h(a)$ <u>implies</u> $h(a) = h(b)$, <u>if and only if</u> $I$ <u>satisfies the following condition</u>: $0 \leq x \leq y \in I$ <u>implies</u> $x \in I$.

PROOF. (a) Trivial.

(b) By (*), $h(a) \leq h(b)$ and $h(b) \leq h(c)$ implies there exist elements $x, y, z \in A$ such that $x \leq y$, $y \leq z$, and $a - x \in I$, $b - y \in I$, $c - z \in I$. Since, $x \leq y$ and $y \leq z$ implies $x \leq z$ the required result follows.

(c) In order to prove (i) assume that $h(a) \leq h(b)$. Then $h(b-a) \geq 0$ as was shown above. Since $h(b-a) = h\{(b + c) - (a + c)\}$ we obtain from $h(b - a) \geq 0$ that $h(b+c) \geq h(a+c)$. Hence, $h(b) + h(c) \geq h(a) + h(c)$ for all $c \in A$.

If $h(a) \geq 0$ and $h(b) \geq 0$, then by (*) there exist elements $x, y \in A$ such that $x \geq 0$, $y \geq 0$ and $a - x \in I$ and $b - y \in I$. Since $x \geq 0$, $y \geq 0$ implies $xy \geq 0$ and $a - x \in I$ and $b - y \in I$ imply $ab - xy \in I$, condition (ii) of Definition 3.4 follows then from (*).

(d) Let $0 \leq x \leq y \in I$. Then $h(y-x) \geq 0$ and $h(x) \geq 0$. Since $h(y) = 0$ we obtain from $h(y-x) = h(y) - h(x) \geq 0$ that $h(x) \leq 0$. Hence, $h(x) = 0$, i.e., $x \in I$. Conversely, assume that $h(a) \leq h(b)$ and $h(a) \geq h(b)$. Then $h(a-b) \geq 0$ and $h(b-a) \geq 0$. Hence, by (*), there exist elements $x, y \in A$ such that $x \geq 0$, $y \geq 0$ and $a-b - x \in I$ and $b-a - y \in I$. We conclude that $x + y \in I$. Since $0 \leq x \leq x + y \in I$, we have $x \in I$ or equivalently $h(a-b) = 0$, i.e., $h(a) = h(b)$. This completes the proof of the theorem.

In summary, if $I$ _is a proper ideal of an ordered ring_ $A$ _such that_ $0 \leq x \leq y \in I$ _implies_ $x \in I$, _then_ $A/I$ _is an ordered ring under_ (*) _and the canonical mapping_ $h$ _of_ $A$ _onto_ $A/I$ _is order preserving_.

REMARK. The reader does well to observe that the canonical mapping $h$ of $A$ onto $A/I$ does not necessarily preserve _strict_ inequalities, i.e., if $a < b$, then it may occur that $h(a) = h(b)$.

The following theorem is an immediate consequence of the preceding theorem.

THEOREM 3.6. Let $A$ be a totally-ordered ring and let $I$ be a proper ideal of $A$ satisfying the condition: $0 < x \leq y \in I$ implies $x \in I$. Then $A/I$ is a totally-ordered ring. Furthermore, if $A$ is archimedean, then $A/I$ is archimedean.

## 4. Construction of a Non-Standard Model for Analysis.

In this section we shall construct a non-standard model for analysis in the form of an ultrapower. The ultrapower is chosen in such a way that it is possible to compare the theory of Laugwitz and Schmieden with the theory of Robinson.

Let $R$ again denote the set of real numbers and let $N$ be the set of all positive integers $n \geq 0$. Let $R^N$ be the set of all mappings of $N$ into $R$. Then we may consider $R^N$ as a ring if we define addition and multiplication as pointwise operations. Furthermore, the ring $R^N$ contains a copy of $R$ in the set of all constant functions. If we introduce an order in $R^N$ in the following way that an element of $R^N$ is non-negative if its range does not contain any negative number, then $R^N$ is the ordered ring employed by Schmieden and Laugwitz. It turns out that $R^N$ is not totally-ordered. It is a non-archimedean ordered ring with divisors of zero. We shall now show how we can remove these unpleasant features.

Let $\mathcal{U}$ be a free-ultrafilter on $N$. In what follows, $\mathcal{U}$ will always be the same free-ultrafilter. We introduce now the following definition (elements of $R^N$ will always be denoted by capitals, $A, B, \ldots, \Omega, \ldots$).

DEFINITION 4.1 (Equality relative to $\mathcal{U}$) Let $A$ and $B$ be two elements of $R^N$. We say that $A$ and $B$ are equal relative to $\mathcal{U}$ if the set $\{n : A(n) = B(n)\} \in \mathcal{U}$ and we denote this by $A =_{\mathcal{U}} B$.

The terminology used in the preceding definition suggests the following theorem.

THEOREM 4.1. The relation $A =_{\mathcal{U}} B$ between the elements of $R^N$ is an equivalence relation.

PROOF. The relation $A =_{\mathcal{U}} B$ is obviously symmetric. In order to prove that it is transitive assume that $A =_{\mathcal{U}} B$ and $B =_{\mathcal{U}} C$. Then $\{n : A(n) = D(n)\} \supseteq \{n : A(n) = B(n)\} \cap \{n : B(n) = D(n)\}$. Since $\mathcal{U}$ is a filter we obtain immediately that $\{n : A(n) = D(n)\} \in \mathcal{U}$, i.e. $A =_{\mathcal{U}} D$.

REMARK. In the theory of Schmieden and Laugwitz equality between the elements of $R^N$ relative to the Fréchet filter of $N$ is studied rather than with respect to a free ultrafilter.

DEFINITION 4.2. The set of the classes of equivalent elements relative to the equivalence relation $=_{\mathcal{U}}$ is denoted by $R^*$. The elements of $R^*$ will be denoted by lower case letters $a, b, c, \ldots$ .

The elements of $a \in R^*$ will be denoted by $A, A', \ldots, A_1, \ldots$ etc.

It is easy to see that $R$ can be embedded in $R^*$ in a natural way. Indeed, if $r \in R$, then the class to which the constant function on $N$ with value $r$ belongs will represent $r$ in $R^*$. It is obvious that this defines an injection (i.e., 1-1 and into) of $R$ into $R^*$. We shall denote the elements of $R$ in $R^*$ in the same way. In order to be able to distinguish between the elements of $R^*$ whether they belong to $R$ or not we shall introduce the following definition.

DEFINITION 4.3 (Standard element). An element of $R^*$ which belongs to $R$ will be called a standard element, all the other elements are called non-standard elements of $R^*$.

We shall now define addition and multiplication in $R^*$.

DEFINITION 4.4 (Addition and multiplication) Let $a, b, c \in R^*$. Then $a + b = c$ if and only if there exist $A \in a$, $B \in b$ and $C \in c$ such that $\{n : A(n) + B(n) = C(n)\} \in \mathcal{U}$, and $ab = c$ if and only if there exist $A \in a$, $B \in b$ and $C \in c$ such that $\{n : A(n)\,B(n) = C(n)\} \in \mathcal{U}$.

In order to justify this definition we have to show that it is independent from the elements $A \in a$, $B \in b$ and $C \in c$. To this end, let $A' \in a$, $B' \in b$, and $C' \in c$, then observing that $\{n : A'(n) + B'(n) = C'(n)\}$ $\supseteq \{n : A(n) = A'(n)\} \cap \{n : B(n) = B'(n)\} \cap \{n : C(n) = C'(n)\} \cap$ $\{n : A(n) + B(n) = C(n)\}$, using the fact that $\mathcal{U}$ is a filter, we obtain that $\{n : A'(n) + B'(n) = C'(n)\} \in \mathcal{U}$. In the same way this is proved for multiplication.

THEOREM 4.2. $R^*$ is a field.

PROOF. It is completely obvious that $R^*$ is a commutative ring with identity 1 (= the class to which the constant function 1 belongs). To complete our proof we have to show that for every $a \in R^*$, $a \neq 0$ there exists an element $b$ such that $ab = 1$. Let $a \in R^*$ and $a \neq 0$. Then

$\{n : A(n) = 0\} \notin \mathcal{U}$. Since $\mathcal{U}$ is an ultrafilter we have by Theorem 2.4 that $\{n : A(n) \neq 0\} \in \mathcal{U}$. Let $B'$ be that element of $R^N$ which is equal to $A$ if $A \neq 0$ and equal to 1, say, if $A = 0$. Then $B = 1/B' \in R^N$ and $ab = 1$, where $b$ is the class to which $B$ belongs.

REMARKS 1. In the proof of Theorem 4.2, that $R^*$ is field, we used for the first time the hypothesis that $\mathcal{U}$ is an ultrafilter.

2. If we consider $R^N$ as a ring in the way indicated earlier, then the set $I$ of all elements $A$ of $R^N$ such that $\{n : A(n) = 0\} \in \mathcal{U}$ is a maximal ideal in $R^N$. Furthermore, it is easy to see that $R^N/I \cong R^*$. Which proves again that $R^*$ is a field (Theorem 3.1).

We shall now introduce an order relation in $R^*$.

DEFINITION 4.5 (Order) Let $a, b \in R^*$. Then we say that $a$ is smaller than $b$ or $b$ is larger than $a$ if and only if there exist $A \in a$, $B \in b$ such that $\{n : A(n) \leq B(n)\} \in \mathcal{U}$. This relation will be denoted by $a \leq b$.

It is easy to see that this definition is independent from the elements $A \in a$ and $B \in b$.

THEOREM 4.3. With the relation $a \leq b$ between the elements of $R^*$ we have that $R^*$ is a totally-ordered field.

PROOF. We shall first prove that $a \leq b$ orders $R^*$. It is obvious that for all $a \in R^*$, $a \leq a$, i.e., the relation $\leq$ is reflexive. In order to prove that it is anti-symmetric assume that $a \leq b$ and $b \leq a$. Then $\{n : A(n) = B(n)\} \supseteq \{n : A(n) \leq B(n)\} \cap \{n : A(n) \geq B(n)\}$. Hence, since $\mathcal{U}$ is a filter $\{n : A(n) = B(n)\} \in \mathcal{U}$ or equivalently $a = b$. In order to prove that it is transitive assume that $a \leq b$ and $b \leq c$. Then, since

$\mathcal{U}$ is a filter and $\{n : C(n) \geq A(n)\} \supseteq \{n : A(n) \leq B(n)\} \cap \{n : B(n) \leq C(n)\}$, we obtain that $a \leq c$. We shall now prove that this relation totally orders $R^*$. To this and let $a, b \in R^*$ and assume not $(a \leq b)$. Since $\mathcal{U}$ is an ultrafilter we must have that $\{n : A(n) > B(n)\} \in \mathcal{U}$. Hence $a \geq b$. Incidentally, we have also shown that $a < b$ if and only if $\{n : A(n) < B(n)\}$. Finally, in order to prove that $R^*$ is a totally ordered field we have to show that $a \geq b$ implies $a + c \geq b + c$ for all $c \in R^*$ and $a \geq 0$ and $b \geq 0$ implies $ab \geq 0$. Since $\{n : A(n) + C(n) \geq B(n) + C(n)\} \supseteq \{n : A(n) \geq B(n)\}$ the former statement follows. The latter statement follows from $\{n : A(n)\, B(n) \geq 0\} \supseteq \{n : A(n) \geq 0\} \cap \{n : B(n) \geq 0\}$. This completes the proof of the theorem.

It is evident from Definitions 4.4 and 4.5 that the embedding of R into $R^*$ is not only an algebraic isomorphism but also an order isomorphism. In the following theorem we will show that R and $R^*$ are not isomorphic and thus $R^*$ is a proper extension of R.

THEOREM 4.4. <u>$R^*$ and R are not isomorphic. In fact, $R^*$ is non-archimedean.</u>

PROOF. Observe that there exists a mapping A of N into R such that A is unbounded on every element of $\mathcal{U}$. Indeed, consider the mapping A of N into R defined by $A(n) = n$ for all $n \in N$. Then A is bounded on the finite subsets of N only but those subsets do not belong to $\mathcal{U}$ since $\mathcal{U}$ is a <u>free</u> ultrafilter. Hence, $|a| > n$ for every $n \in N$, where $A \in a$. Thus we have shown that $R^*$ is non-archimedean. Finally, R and $R^*$ cannot be isomorphic as fields. Indeed, every isomorphism of R to a totally ordered field is <u>order-preserving</u>, since every positive element in R has a square noot. The order in R is thus completely determined by the algebraic structure.

REMARK. The following question seems natural: How many different systems $R^*$ may one obtain by choosing different free-ultrafilters on $N$? The answer is that under the continuum hypothesis one can show that the systems $R^*$ are all isomorphic. This follows from results contained in [9].

## 5. Infinitesimals, Finite and Infinitely Large Numbers.

The fact that $R^*$ is non-archimedean suggests the following definition.

DEFINITION 5.1 (Infinitesimals, finite and infintely large numbers). A number $a \in R^*$ is called infinitely large if there exists an element $A \in a$ such that $A$ is unbounded on every element of $\mathcal{U}$. A number $a \in R^*$ is called infinitely small or an infinitesimal if it is the reciprical of an infinite large number or zero. Numbers which are not infinitely large are called finite.

These numbers can be characterized in a slightly different way as follows:

THEOREM 5.1 (i) A number $a \in R^*$ is infinitely large if and only if $|a| > n$ for all $n \geq 1$.

(ii) A number $a \in R^*$ is an infinitesimal if and only if $|a| < 1/n$ for all $n \geq 1$.

(iii) A number $a \in R^*$ is finite if and only if there exists a number $n \in N$ such that $|a| < n$.

PROOF. This theorem is an immediate consequence of Definition 5.1.

The reader may also observe that (i) $a \in R^*$ is infinitely large if and only if $|a| > r$ for all standard numbers $r$, (ii) $a \in R^*$ is infinitesimal if and only if $|a| < r$ for all $r > 0$ and $r$ standard and (iii) $a \in R^*$ is finite if and only if there exists a standard number $r$ such that $|a| < r$.

NOTATION. *The set of all finite numbers will be denoted by* $M_0$ *and the set of all infinitesimals will be denoted by* $M_1$.

It is easv to see that $M_0$ *is a totally-ordered archimedean integral domain. Furthermore,* $R^*$ *is isomorphic to the ring of quotients of* $M_0$.

THEOREM 5.2. $M_1$ *is a maximal ideal in* $M_0$ *and* $M_0/M_1$ *is isomorphic to* $R$.

PROOF. We shall first prove that $M_1$ is an ideal in $M_0$. For this purpose, let $h$ and $k$ be infinitesimals. Then for every $r \in R$, $r > 0$ we have $|h| < r$ and $|k| < r$. Hence, $|h - k| < 2r$ for all $r \in R$ and $r > 0$, i.e., $h - k \in M_1$. Furthermore, if $h \in M_1$ and $a \in M_0$, then $|ah| < rs$ for all $r \in R$ and where $s \in R$ is such that $|a| < s$. Hence, $ah \in M_1$. This shows that $M_1$ is an ideal. In order to show that $M_1$ is a maximal ideal observe that if $a \in M_0$ but $a \notin M$, then $a^{-1} \in M_0$. Hence, by Theorem 3.3, $M_1$ is a maximal ideal. Furthermore, if $0 \leq a \leq h \in M_1$, then $0 \leq a < r$ for all $r \in R$ and $r > 0$. Hence, $a \in M_1$, i.e., $M_1$ is a maximal ideal which satisfies the condition of (d) of Theorem 3.5. We conclude from Theorems 3.3 and 3.6 that $M_0/M_1$ is a totally-ordered field which is archimedean since $M_0$ is archimedean. Hence, by Theorem 3.4, $M_0/M_1$ is isomorphic to a subfield of $R$. Since $M_0/M_1$ contains a copy of $R$, we obtain finally that $M_0/M_1$ is isomorphic to $R$. This completes the proof of the theorem.

The homomorphism of $M_0$ onto $R$ with kernel $M_1$ plays a very important role in non-standard analysis. We shall introduce therefore the following definition.

DEFINITION 5.2 (Standard part of a finite number) The homomorphism of $M_0$ onto $R$ with kernel $M_1$ will be denoted by st. Furthermore, for every $a \in M_0$ the unique standard number $st(a)$ of $a$ will be called the standard part of $a$.

We shall write $a =_1 b$ if $a - b \in M_1$, i.e., if the difference between $a$ and $b$ is infinitesimal. In that case, we say that $a$ is infinitely close to $b$. From Theorem 5.2 it follows that every finite number is infinitely close to a unique standard number, namely its standard part.

We shall conclude this section with the following list of properties of the homomorphism st of $M_0$ onto $R$ with kernel $M_1$ which are frequently used in non-standard analysis.

(i) For all $a, b \in M_0$ we have $st(a + b) = st(a) + st(b)$ and $st(ab) = st(a)\ st(b)$.

(ii) For all $a, b \in M_0$ we have that $a \leq b$ implies $st(a) \leq st(b)$.

(iii) For all $a, b \in M_0$ we have that $st(\max(a,b)) = \max(st(a), st(b))$ and $st(\min(a,b)) = \min(st(a), st(b))$. In particular, $st(|a|) = |st(a)|$ for all $a \in M_0$.

(iv) We have $st(a) = 0$ if and only if $a \in M_1$. Hence, if $s,t \in R$ and $s < t$, then $s + h < t + k$ for all infinitesimals $h, k \in M_1$. This shows, in pafticular, that (ii) cannot be impeoved to $a < b$ implies $st(a) < st(b)$. Indeed, $a < b$ implies $st(a) = st(b)$ if and only if $a - b \in M_1$.

(v) If $a \in M_0$, then $st(a) \geq 0$ if and only if $a =_1 |a|$

(vi) If $r \in R$, then $st(r) = r$.

## 6. The Interval Topology of $R^*$.

For the terminology which is not explained in this section the reader is referred to: J. L. Kelley, General Topology, van Nostrand (1955) and N. Bourbaki, Livre III, Topologie Générale, Chap. I. Structures topologiques, third edition, Hermann 1142 (1961).

In the preceding section we have discussed the algebraic properties of $R^*$. In this section we shall deal briefly with the problem of making $R^*$ in a topological space. We begin by introducing the important concept of an interval.

DEFINITION 6.1 (Interval). A non-empty subset $I$ of $R^*$ is called an interval if $a, b \in S$ and $a < b$ implies $x \in S$ for all $a < x < b$.

Of course the following sets are examples of intervals:
$\{x : a < x < b\}$, $\{x : a < x \leq b\}$, $\{x : a \leq x < b\}$, $\{x : a \leq x \leq b\}$, $\{x : x > a\}$, $\{x : x < b\}$, etc.... . Furthermore, the intersection of a finite collection of intervals is an interval. The union of two intervals is in general not an interval. If, however, their intersection is non-empty, then their union is an interval.

We may make $R^*$ into a topological space by taking as a subbase for its open sets the family of all rays $\{x : x > a\}$ and $\{x : x < b\}$. This topology is called the interval topology of $R^*$. In this topology all intervals of the type $\{x : a < x < b\}$ are open. When an interval is an open set, it is called an open interval. Since the open intervals of $R^*$ form a base, the topology is called the interval topology. It is easy to see that $R^*$ is a Hausdorff space. Furthermore, the operations of addition and multiplication are continuous. Thus, $R^*$ is a topological field in its interval topology.

The topology of $R$ is of course its interval topology. The reader should notice, however, that the interval topology of $R^*$ induces on $R$ the discrete topology. Indeed, if $r \in R$ and $0 < h \in M_1$, then $|a - r| < h$ and $a \neq r$ implies $a \notin R$. Hence, this neighborhood of $r$ intersects $R$ in $r$. In other words, the embedding of $R$ in $R^*$ is not topological. It is a discontinuous embedding.

THEOREM 6.1. Every open set of $R^*$ is expressible in a unique way as a union of disjoint maximal open intervals.

PROOF. Let $S$ be an open subset of $R$ and let $x \in S$. Let the set $O$ of all open intervals $I$ such that $x \in I$ and $I \subseteq S$ be ordered by inclusion. Since $S$ is open, $O$ is non-empty. It is easy to see that the union of the elements of $O$ is an element of $O$. This interval, which is a maximal open interval contained in $S$ and containing $x$, is denoted by $I_x$. Then the family of open intervals $\{I_x : x \in S\}$ has the required properties. The uniqueness follows easily.

Another fact, worth noticing, is that $R^*$ is disconnected. Indeed, it is, for instance, easy to see that set of all infinitesimals $M_1$ in $R^*$ is both open and closed. Also, the set of all infinitely large numbers is both open and closed. The same holds for the set of all finite elements $M_0$. The fact that $R^*$ is disconnected in its interval topology can also be expressed in the following form: $R^*$ is not Dedekind complete (An ordered set is called Dedekind complete if every non-empty subset which is bounded above has a least upper bound). Indeed, if $a$ is a least upper bound of $S \subseteq R^*$, then $a$ is obviously also in the closure of $S$. The non-empty and bounded set $M_1$ has no least upper bound in $R^*$.

We conclude this section with the remark that it can also be shown that $R^*$ is a normal space in its interval topology.

## 7. Non-Standard Extensions

In the preceding sections, in particular in section 5, we have given a number of well-known facts and notions concerning all totally-ordered fields which are extensions of $R$. In this section we shall make use of the fact that $R^*$ is an ultrapower. We shall see that this fact will enable us to reformulate the notions and procedures of classical analysis in $R^*$.

The following definition is essential.

DEFINITION 7.1 (*Non-standard extension of a standard set*) *Let* $S$ *be a subset of* $R$. *The set of all* $a \in R^*$ *for which there exists an element* $A \in a$ *such that* $\{n : A(n) \in S\} \in \mathcal{U}$ *is called the non-standard extension of* $S$ *and is denoted by* $S^*$.

It is obvious that $(\exists A)(A \in a$ and $\{n : A(n) \in S\} \in \mathcal{U}) \Rightarrow (\forall A)(A \in a \Rightarrow \{n : A(n) \in S\} \in \mathcal{U})$. Hence, Definition 7.1 can be given in symbols as follows: $a \in S^* \Leftrightarrow (\forall A)(A \in a \Rightarrow \{n : A(n) \in S\} \in \mathcal{U})$.

THEOREM 7.1 *The following results hold:*

(i) *The non-standard extension of the empty subset of* $R$ *is the empty subset of* $R^*$, *i.e.,* $\phi^* = \phi$.

(ii) *For every subset* $S$ *of* $R$, $S \subseteq S^*$.

(iii) *If* $S_1$ *and* $S_2$ *are subsets of* $R$, *then* $S_1 \subseteq S_2$ *implies* $S_1^* \subseteq S_2^*$.

(iv) *If* $S_1$ *and* $S_2$ *are subsets of* $R$, *then* $(S_1 \cap S_2)^* = S_1^* \cap S_2^*$ *and* $(S_1 \cup S_2)^* = S_1^* \cup S_2^*$. *In particular* $(C_R(S))^* = C_{R^*}(S^*)$ *for every subset* $S$ *of* $R$, *where* $C_R(S)$ *and* $C_{R^*}(S^*)$ *are the complements of* $S$ *in* $R$ *and* $S^*$ *in* $R^*$ *respectively.*

(v) *If* $S$ *is a subset of* $R$ *and* $a \in S^*$, *then* $a \in S$ *if and only if* $a$ *is a standard number.*

PROOF. (i) If $S = \emptyset$, then $\{n : A(n) \in S\} = \emptyset$. Hence, $\emptyset^* = \emptyset$.

(ii) If $r \in S$, then $\{n : r \in S\} = N \in \mathcal{U}$, i.e., $r \in S^*$.

(iii) If $S_1 \subseteq S_2$, then $\{n : A(n) \in S_1\} \subseteq \{n : A(n) \in S_2\}$. Since $a \in S_1^*$ implies $\{n : A(n) \in S_1\} \in \mathcal{U}$ for all $A \in a$ we obtain, using the fact that $\mathcal{U}$ is a filter, that $\{n : A(n) \in S_2\} \in \mathcal{U}$ for all $A \in a$. Hence, $a \in S_1^*$ implies $a \in S_2^*$, i.e., $S_1^* \subseteq S_2^*$.

(iv) From $S_1 \cap S_2 \subseteq S_1$ and $S_1 \cap S_2 \subseteq S_2$ it follows immediately, using (iii), that $(S_1 \cap S_2)^* \subseteq S_1^* \cap S_2^*$. Let $a \in S_1^* \cap S_2^*$. Then $\{n : A(n) \in S_1\} \in \mathcal{U}$ and $\{n : A(n) \in S_2\} \in \mathcal{U}$. Hence, since $\mathcal{U}$ is a filter, $\{n : A(n) \in S_1\} \cap \{n : A(n) \in S_2\} \in \mathcal{U}$. But $\{n : A(n) \in S_1\} \cap \{n: A(n) \in S_2\} = \{n : A(n) \in S_1 \cap S_2\}$. We conclude that $a \in (S_1 \cap S_2)^*$. From $S_1 \subseteq S_1 \cup S_2$ and $S_2 \subseteq S_1 \cup S_2$ it follows, using (iii), that $S_1^* \cup S_2^* \subseteq (S_1 \cup S_2)^*$. Let $a \in (S_1 \cup S_2)^*$. Then $\{n : A(n) \in S_1 \cup S_2\} \in \mathcal{U}$ for all $a \in a$. Since $\{n : A(n) \in S_1 \cup S_2\} = \{n : A(n) \in S_1\} \cup \{n : A(n) \in S_2\}$ and $\mathcal{U}$ is an ultrafilter, Theorem 2.2 implies that either $\{n : A(n) \in S_1\} \in \mathcal{U}$ or $\{n : A(n) \in S_2\} \in \mathcal{U}$. Hence, $a \in (S_1 \cup S_2)^*$ implies $a \in S_1^* \cup S_2^*$. We conclude that $(S_1 \cup S_2)^* = S_1^* \cup S_2^*$. Finally, from $S \cup C_R(S) = R$ and $S \cap C_R(S) = \emptyset$ it follows that $S^* \cup (C_R(S))^* = R^*$ and $S^* \cap (C_R(S))^* = \emptyset$, i.e., $(C_R(S))^* = C_{R^*}(S^*)$.

(v) We have only to show that $a \in S^*$ and $a$ is standard implies that $a \in S$. But this follows immediately from the fact that in that case $\{n : a \in S\} = N \neq \emptyset$. This completes the proof of the theorem.

REMARKS 1. As an immediate corollary to (iv) of the preceding theorem we have that <u>if</u> $S_1, \ldots, S_n$ <u>are subsets of</u> $R$, <u>then</u> $(\bigcup_{i=1}^{n} S_i)^* = \bigcup_{i=1}^{n} S_i^*$ <u>and</u> $(\bigcap_{i=1}^{n} S_i)^* = \bigcap_{i=1}^{n} S_i^*$. Indeed, induction on $n$ will prove it.

2. From (i), (iii) and (iv) of Theorem 7.1 it follows that the mapping $S \to S^*$ of the Boolean algebra of all subsets of $R$ into the Boolean algebra of all subsets of $R^*$ is a <u>Boolean homomorphism</u>, i.e., it preserves the Boolean operations. Since $S \neq \emptyset$ implies that $S^* \neq \emptyset$ ( this follows from (ii) of Theorem 7.1) we see that this mapping is actually an <u>isomorphism</u> of the Boolean algebra of all subsets of $R$ onto a <u>proper</u>-subalgebra of the Boolean algebra of all subsets of $R^*$. The reader is advised to check that the isomorphism is indeed <u>not an isomorphism onto</u> the algebra of all subsets of $R^*$.

In the following three theorems we shall prove some more useful properties of the non-standard extension of a standard set.

THEOREM 7.2 <u>If</u> $S$ <u>is a finite subset of</u> $R$, <u>then</u> $S = S^*$.

PROOF. It follows immediately from property (iv) of Theorem 7.1 that if $S_1, \ldots, S_n$ are subsets of $R$, then $\bigcup_{i=1}^{n} S_i^* = (\bigcup_{i=1}^{n} S_i)^*$. Indeed, induction on $n$ will prove it. Hence, if $S \subseteq R$ is finite, i.e., $S = \{x_i : 1 \leq i \leq n\}$, then in order to prove the theorem we have to show only that $\{x\}^* = \{x\}$ for all $x \in R$. To this end assume that $a \in \{x\}^*$. Then $\{n : A(n) = x\} \in \mathcal{U}$, where $A \in a$. Hence, by Definition 4.2, $a = x$, i.e., $\{x\}^* = \{x\}$.

In the following theorem we shall formulate a result which is an immediate consequence of Theorem 7.2, and (iv) of Theorem 7.1. It is given here because of the important role it plays in the theory of limits.

THEOREM 7.3. <u>If</u> $S \subseteq R$ <u>and</u> $T$ <u>is a finite subset of</u> $S$, <u>then</u> $(S - T)^* = S^* - T$.

PROOF. If we write $S = (S - T) \cup T$ and use (iv) of Theorem 7.1 we obtain that $S^* = (S - T)^* \cup T^*$. Since $T$ is finite, $T = T^*$ by the preceding theorem. Hence, $S^* = (S - T)^* \cup T$. We conclude that $S^* - T = (S - T)^*$, which completes the proof of the theorem.

We shall now prove a theorem which plays an important role in a non-standard proof we shall give for the Bolzano-Weierstrass theorem.

THEOREM 7.4. <u>A subset</u> $S$ <u>of</u> $R$ <u>is infinite if and only if</u> $S \neq S^*$.

PROOF. It follows from Theorem 7.2 that $S \neq S^*$ implies that $S$ is infinite. To prove the converse, assume that $S$ is an infinite subset of $R$. Then there exists an injection $A$ of $N$ into $S$. Hence, if $a$ is the class to which $A$ belongs, then $a \in S^*$. If we now assume that $S = S^*$, then $a \in S$, i.e., $a$ is standard. Since $A$ is an injection, there exists one and only one index, say $n_0$, such that $A(n_0) = a$. By definition, $\{n_0\} = \{n : A(n) = a\} \in \mathcal{U}$. This contradicts the fact that $\mathcal{U}$ is a free ultrafilter and completes the proof of the theorem.

The preceding theorems allows us to show that the isomorphism $S \to S^*$ does <u>not preserve infinite intersections and unions</u>. If $\{S_i : i \in I\}$ is an arbitrary, not necessarily countable, family of subsets of $R$, then (iii) of Theorem 7.1 implies that $(\bigcup_{i \in I} S_i)^* \supseteq \bigcup_{i \in I} S_i^*$ and $(\bigcap_{i \in I} S_i)^* \subseteq \bigcap_{i \in I} S_i^*$.

Equality may not hold if $I$ is infinite as the following examples will show. Let $\{\{x_n\} : n \in N\}$ be a countable family of unit sets such that $x_n \neq x_m$ as $n \neq m$. Then, by Theorem 7.2, $\bigcup_{n\in N} \{x_n\}^* = \bigcup_{n\in N} \{x_n\}$. But Theorem 7.4 implies that $\bigcup_{n\in N} \{x_n\} \neq (\bigcup_{n\in N} \{x_n\})^*$. Next, let $N_n$ denote the set of all natural numbers $m \geq n$. Then $\bigcap_{n\in N} N_n = \phi$. Hence, $(\bigcap_{n\in N} N_n)^* = \phi$. But Theorem 7.3 implies that for all $n \in N$, $N_n^* - N_n = N^* - N$. Since, by Theorem 7.4, $N^* - N \neq \phi$, we have $\bigcap_{n\in N} N_n^* \neq \phi$. Thus $(\bigcap_{n\in N} N_n)^* = \phi \neq \bigcap_{n\in N} N_n^*$ $(= N^* - N)$.

REMARK. It is of importance for the reader to be aware of the fact that the non-standard extension $S^*$ of a subset $S$ can be looked upon as the set defined in $R^*$ by the singularly relation (i.e. a relation with one free variable) in $R$ which defines the set $S$ in $R$. This method of enriching our vocabulary is an important tool in the theory of non-standard analysis. It is an immediate consequence of the construction of $R^*$ as an ultrapower. We may carry this idea a little further. If $S$ is a subset of $R$ and $\Phi$ is some singularly relation which holds for all the elements of $S$, then it holds in $S^*$ in the following sense: $a \in S^*$, then $\{n : \Phi(A(n))\} \in \mathcal{U}$, where $A \in a$. The sentence or formula which expresses this in $R^*$ can be obtained immediately from the sentence $\Phi$ in $R$ by replacing the individual statements by their extension in $R^*$ provided that they have been extended already. Such statements are e.g. addition, multiplication, order and statements like $x \in S$. We shall illustrate this method by means of an example. Let $I$ be an interval in $R$. Then its non-standard extension $I^*$ is an interval in $R^*$. This can be proved directly as follows: Let $a$, $b \in I^*$ and let $x \in R^*$ be such that $a < x < b$. Then $\{n : A(n) < X(n) < B(n)\} \in \mathcal{U}$. Since $\{n : A(n) \in I\} \in \mathcal{U}$ and $\{n : B(n) \in I\} \in \mathcal{U}$, we obtain immediately that $\{n : X(n) \in I\} \in \mathcal{U}$. Hence, $x \in I^*$. Let us now look at it from the following point of view. The sentence in $R$ which expresses that $I$ is an interval reads: $(\forall a)(\forall b)(\forall x)(a \in I,\ b \in I,\ x \in R$ and $a < x < b \Rightarrow x \in I)$. Hence, we have $(\forall a)(\forall b)(\forall x)(a \in I^*,\ b \in I^*,\ x \in R^*$ and $a < x < b \Rightarrow x \in I^*)$, i.e., $I^*$ is an interval.

Similarly, we can show that the non-standard extension of an open subset of $R$ is an open subset of $R^*$ in its interval topology. The reader is advised to check this.

Recall that a non-empty subset $S$ of $R$ is said to be <u>bounded</u> if there exists an element $r \in R$ such that $|x| < r$ for all $x \in S$. It is said to be <u>bounded above</u> (resp. <u>bounded below</u>) if there exists an element $r \in R$ such that $x \leq r$ (resp. $x \geq r$) for all $x \in S$. In non-standard analysis this can be expressed as follows:

THEOREM 7.5. <u>A non-empty subset</u> $S$ <u>of</u> $R$ <u>is bounded if and only if</u> $S^* \subseteq M_o$. <u>It is bounded above</u> (resp. <u>bounded below</u>) <u>if and only if</u> $S^*$ <u>has no infinitely large positive numbers</u> (resp. <u>infinitely large negative numbers</u>).

PROOF. If $S$ is bounded, then there exists an element $r \in R$ such that $|x| < r$ for all $x \in S$. Hence, $|x| < r$ for all $x \in S^*$, i.e. $S^* \subseteq M_o$. In order to prove the converse, we shall assume that $S$ is not bounded but that $S^* \subseteq M_o$. The statement that $S$ is not bounded can be expressed by the sentence: $(\forall r)(r \in R \Rightarrow (\exists x)(x \in S$ and $|x| > |r|))$. Hence, $(\forall r)(r \in R^* \Rightarrow (\exists x)(x \in S^*$ and $|x| > |r|))$. By taking then for $r$ an infinitely large number we obtain that $S^*$ has an infinitely large element. This contradicts the assumption that $S^* \subseteq M_o$.

If $S$ is bounded above, then there exists an element $r \in R$ such that $x < r$ for $x \in S$. Hence, $x < r$ for all $x \in S^*$ which shows that $S^*$ has no infinitely large positive numbers. Conversely, if $S^*$ has no infinitely large positive numbers, then $S$ is bounded above. Indeed, if $S$ is <u>not</u> bounded above, then for every $r \in R$ there exists an element $x \in S$ such that $x > r$. Hence, for every $r \in R^*$ there exists an element $x \in S^*$ such that $x > r$. If we take for $r$ an infinitely large positive number we obtain that $S^*$ contains an infinttely large positive number. This contradicts the assumption. The proof that $S$ is bounded below if and only if $S^*$ has no infinitely large negative numbers is similar to the preceding proof.

REMARK. If $S \subseteq R$ and $S^* \subseteq M_0$, then there exists already an element $r \in R$ such that $|x| < r$ for all $x \in S^*$. Hence, in particular, there does not exist a set $S \subseteq R$ such that $S^* = M_0$.

EXAMPLES 1. (<u>A non-standard model for</u> $N$) The set of all positive integers $N(n \geq 0)$ defines a subset $N^*$ in $R^*$. Then $N^*$ constitutes a non-standard model for $N$ in the sense of Skolem, i.e., with respect to all properties of $N$ that are formulated in the lower predicate calculus in terms of some given set of number theoretic relations or functions. Since $N$ is infinite, it follows from Theorem 7.3 that $N^* \neq N$. In addition to this, the following is true.

THEOREM 7.6. <u>We have</u> $N^* \cap M_0 = N$.

PROOF. If $a \in N^*$ and $a$ is finite, then it is evident that if $A \in a$, $A$ can take on at most a <u>finite</u> number of different values of $N$, say $n_1, \ldots, n_p$. Then $\bigcup_{i=1}^{p} \{n : A(n) = n_i\} \in \mathcal{U}$. Hence, by Theorem 2.3, there exists precisely one index $i$ such that $\{n : A(n) = n_i\} \in \mathcal{U}$ (use the fact that the sets $\{n : A(n) = n_i\}$, $i=1,2,\ldots,p$, are disjoint). Hence, $a = n_i$, which completes the proof of the theorem.

As an immediate consequence of this theorem we have that the numbers in the non-empty set $N^* - N$ are all infinitely large. This justifies the following definition.

DEFINITION 7.2 (Infinitely large natural numbers). The elements of $N^* - N$ are called the infinitely large natural numbers and will be denoted by $\omega, \mu, \ldots$ with and without super or subscripts.

An infinitely large natural number is called even if it is divisible by 2. If it is not even, then it is called odd. Thus $\omega \in N^* - N$ is even if and only if there exists (and hence it holds for all) an element $\Omega \in \omega$ such that $\{n : \Omega(n)$ is divisible by $2\} \in \mathcal{U}$. If temporarily we denote the set of even natural numbers by $E$ and the set of all odd natural numbers by $O$, then $E^* - E$ is the set of all infinitely large natural numbers which are even and $O^* - O$ is the set of all infinitely large natural numbers which are odd. Furthermore, $N^* - N = (E^* - E) \cup (O^* - O)$.

If $P$ denotes temporarily the set of all prime numbers, then the elements of the set $P^* - P$ are called the infinitely large prime numbers (observe that $P^* - P$ is not empty since $P$ is an infinite subset of $N$). The elements of $P^* - P$ can be characterized as follows: $\omega \in P^* - P$ if and only if there exists an element $\Omega \in \omega$ such that $\{n : \Omega(n)$ is prime$\} \in \mathcal{U}$.

The arithmetic functions in $N$ extend naturally to $N^*$, e.g. if $\omega \in N^* - N$, then $\omega!$ denotes the class to which the element $\Omega(n)!$ belongs, where $\Omega \in \omega$. We have also in this case that $(\omega + 1)! = (\omega + 1)\,\omega!$.

If $r > 0$ is a standard number, then $[r]$ usually denotes the largest natural number contained in $r$. This function extends naturally. Indeed if $a \in R^*$, then $[a]$ is that element of $N^*$ such that if $A \in a$, then $[A] \in [a]$, where $[A]$ denotes the function $[A(n)]$. If $a$ is infinitely large, then $[a] \in N^* - N$. We may also call $[a]$ the largest natural number less than $a$.

2. (<u>Non-standard model for</u> $Q$) The field of rationals $Q$ defines a set $Q^* \neq Q$ in $R^*$. $Q^*$ is a non-standard model for $Q$. Furthermore, $Q^*$ is a totally-ordered non-archimedean subfield of $R^*$. The following argument is another illustration of the method explained in the Remark following Theorem 7.3. It is well-known that $Q$ is dense in $R$, i.e., if $a, b \in R$ and $a < b$, then there exists a rational $r$ such that $a < r < b$. Then this holds in $R^*$ also. Indeed, let $a, b \in R^*$ and assume that $a < b$. Then $E = \{n : A(n) < B(n)\} \in \mathcal{U}$. ..

For every $n \in N$ there exists a rational $P(n) \in Q$ such that $A(n) < P(n) < B(n)$. Define $P(n)=0$ outside $E$ and let $p \in Q^*$ be that element of $Q^*$ such that $P \in p^*$. Then obviously $a < p < b$. In fact, this shows that the sentence which expresses that $Q$ is dense in $R$ : $(\forall a)(\forall b)(a,b \in R$ and $a < b \Rightarrow (\exists r)(r \in Q$ and $a < r < b))$ expresses the same property for $Q^*$ in $R^*$. One has to replace only the separate statements $a,b \in R$, $r \in Q$, $a < r < b$ and $a < r < b$, by the corresponding statements in $R^*$.

Another example is the following. If $z$ is a standard transcendental number, then $(\forall \varepsilon)(\varepsilon \in R$ and $\varepsilon > 0 \Rightarrow (\exists r)(r \in Q$ and $|z-r| < \varepsilon))$. Hence, $(\forall \varepsilon)(\varepsilon \in R^*$ and $\varepsilon > 0 \Rightarrow (\exists r)(r \in Q^*$ and $|z-r| < \varepsilon))$. By taking for $\varepsilon$ a positive infinitesimal in the latter statement we obtain that <u>every standard transcendental number is infinitely close to some element of</u> $Q^*$.

3. Let $R_+$ denote the set of all strictly positive real numbers, i.e., $R_+ = \{x : x \in R \text{ and } x > 0\}$. Then $R^*_+ = (R^*)_+$ is the set of all strictly positive numbers of $R^*$.

We shall now discuss the non-standard extension of binary relations in R. Thus enriching our vocabulary further.

DEFINITION 7.3 (Non-standard extension of binary relations) Let $\Phi$ be a binary relation in R, i.e., $\Phi \subseteq R \times R$. Then the relation $(\exists A)(\exists B)(A \in a \text{ and } B \in b \text{ and } \{n : (A(n), B(n)) \in \Phi\} \in \mathcal{U})$ between the elements a,b of $R^*$ is called the non-standard extension of $\Phi$ and will be denoted by $\Phi^*$.

It is evident that $\Phi^* \subseteq R^* \times R^*$ and hence defines a binary relation in $R^*$. Furthermore, using the fact that $\mathcal{U}$ is a filter, it is obvious that $(\exists A)(\exists B)(A \in a \text{ and } B \in b \text{ and } \{n : (A(n), B(n)) \in \Phi\} \in \mathcal{U}) \Rightarrow (\forall A)(\forall B)(A \in a \text{ and } B \in b \Rightarrow \{n : (A(n), B(n)) \in \Phi\} \in \mathcal{U})$.

If we denote the domain of $\Phi$ by $\Delta_1 \Phi$, i.e., $\Delta_1 \Phi = \{x : (\exists y)((x,y) \in \Phi)\}$, and if we denote the range of $\Phi$ by $\Delta_2 \Phi$, i.e., $\Delta_2 \Phi = \{y : (\exists x)((x,y) \in \Phi)\}$, then the following result holds.

THEOREM 7.7. For every binary relation $\Phi \subseteq R \times R$ we have $(\Delta_1 \Phi)^* = \Delta_1 \Phi^*$ and $(\Delta_2 \Phi)^* = \Delta_2 \Phi^*$.

PROOF. Because of the symmetry it is sufficient to show that $(\Delta_1 \Phi)^* = \Delta_1 \Phi^*$. To this end let $a \in (\Delta_1 \Phi)^*$. Then $\{n : A(n) \in \Delta_1 \Phi\} \in \mathcal{U}$ or equivalently $\{n : (\exists B(n))((A(n), B(n)) \in \Phi)\} \in \mathcal{U}$. Hence $(\exists b)((a,b) \in \Phi^*)$ i.e., $a \in \Delta_1 \Phi^*$. Conversely, $a \in \Delta_1 \Phi^*$ implies $(\exists b)((a,b) \in \Phi^*)$ or equivalently $\{n : (\exists B(n))((A(n),B(n)) \in \Phi)\} \in \mathcal{U}$. Hence, $\{n : A(n) \in \Delta_1 \Phi\} \in \mathcal{U}$, i.e., $a \in (\Delta_1 \Phi)^*$.

It follows immediately from Definition 7.3 that whatever property $\Phi$ may possess, which is expressible in the language of R, other than being a relation is also possessed by $\Phi^*$. For instance, if $\Phi$ is symmetric, then $\Phi^*$ is symmetric; if $\Phi$ is an equivalence relation, then $\Phi^*$ is an equivalence relation. The latter statement may be illustrated by the following example. Let the relation $\Phi$ be defined as follows: $(x,y) \in \Phi \Leftrightarrow (\exists r)(r \in Q$ and $x-y = r)$. Then $\Phi$ is an equivalence relation. Its non-standard extension takes the form: $(x,y) \in \Phi^* \Leftrightarrow (\exists r)(r \in Q^*$ and $x - y = r)$, which is obviously an equivalence relation.

An important property of a relation $\Phi$ is that of being single-valued, i.e., $(\forall x)(\forall y)(\forall z)((x,y) \in \Phi$ and $(x,z) \in \Phi \Rightarrow y = z)$ or in other words $\Phi$ is a function. We obtain then immediately that the <u>non-standard extension of a function is a function</u>. In particular, if $f$ is a mapping of $A \subset R$ into $B \subset R$, then its non-standard extension $f^*$ is a mapping of $A^*$ into $B^*$. In that case, Theorem 7.7 implies that $(f(A))^* = f^*(A^*)$. This procedure allows us in a very simple way to extend the elementary functions to $R^*$. We shall illustrate that by means of a number of examples.

EXAMPLES. 1. (<u>The exponential function</u>) The exponential function $e^x$ is a bijection of $R$ onto $R_+$. Hence, its non-standard extension $\overset{*}{e}{}^x$ is a bijection of $R^*$ onto $R^*_+$. Its functional equation $e^{x+y} = e^x e^y$ carries over, i.e., $\overset{*}{e}{}^{x+y} = \overset{*}{e}{}^x \overset{*}{e}{}^y$. If $h$ is infinitesimal, then $\overset{*}{e}{}^h =_1 1$. Indeed, if $h$ is infinitesimal and $H \in h$, then for every $\varepsilon > 0$ we have $\{n : |e^{H(n)} - 1| < \varepsilon\} \in \mathcal{U}$.

THEOREM 7.8. <u>The non-standard exponential function</u> $\overset{*}{e}^{x}$ <u>maps</u> $M_o$ <u>into</u> $M_o$ <u>and we have</u> $st(\overset{*}{e}^{a}) = e^{st(a)}$ <u>for all</u> $a \in M_o$.

PROOF. Let $a \in M_o$. Then $a = st(a) + h$, where $h \in M_1$. Hence $\overset{*}{e}^{a} = \overset{*}{e}^{(st(a)+h)} = e^{st(a)}\,\overset{*}{e}^{h}$. We conclude that $\overset{*}{e}^{a} \in M_o$ and $st(\overset{*}{e}^{a}) = e^{st(a)}\,st(\overset{*}{e}^{h}) = e^{st(a)}$.

2. (<u>The logarithmic function</u>). The logarithmic function $\log x$ is a bijection of $R_+$ onto $R$. Hence, its non-standard extension $\log^* x$ is a bijection of $R^*_+$ onto $R^*$. The functional equation of $\log x$ carries over. Thus $\log^* xy = \log^* x + \log^* y$ for all $x,y \in R^*_+$. Since $e^{\log x} = x$ for all $x \in R_+$ we have also $\overset{*}{e}^{\log^* x} = x$ for all $x \in R^*_+$. Hence the functions $e^*$ and $\log^*$ are inverses of each other. If $a \in M_o$ but $a \notin M_1$, then $st(\overset{*}{e}^{\log^* a}) = st(a)$. But $st(\overset{*}{e}^{\log^* a}) = e^{st(\log^* a)}$. Hence $e^{st(\log^* a)} = st(a) = e^{\log st(a)}$, i.e., $st(\log^* a) = \log st(a)$. From $\overset{*}{e}^{(-\log^* x)} = 1/x$ it follows in particular, using Theorem 7.6, that if $h$ is infinitesimal, then $-\log^* h$ is an infinitely large positive number.

3. (<u>Trigonometric functions</u>) The trigonometric functions $\sin x$ and $\cos x$ are bounded $2\pi$-periodic functions. Hence, there non-standard extensions $\sin^* x$ and $\cos^* x$ have the same properties. Furthermore, $(\sin^* x)^2 + (\cos^* x)^2 = 1$ for all $x \in R^*$. The other well-known formulas for sin and cos hold for $\sin^*$ and $\cos^*$ as well. If $h$ is infinitesimal, then $\sin^* h =_1 h$. Indeed, from $|\sin x| \leq |x|$ it follows that $\sin^* h$ is infinitesimal for infinitesimal $h$. Hence, $\sin^* h - h =_1 0$. From $\sin^* h =_1 h$ it follows then, using the relation $(\cos^* h)^2 = 1 - (\sin^* h)^2$, that $\cos^* h =_1 1$. Now if $a \in M_o$, then $\sin^* a = \sin^*(st(a) + h) = \sin(st(a))\cos^* h + \cos(st(a))\sin^* h$. Hence, $st(\sin^* a) = \sin(st(a))$. The same result

holds for the cosine function. If $0 < x < \frac{\pi}{2}$, then $\sin x < x < \tan x$ or equivalently $\cos x < \frac{\sin x}{x} < 1$. Hence, if $h > 0$ and $h \in M_1$, then $\overset{*}{\cos} h < \frac{\overset{*}{\sin} h}{h} < 1$. Since $\overset{*}{\cos} h =_1 1$ we obtain that $\frac{\overset{*}{\sin} h}{h} =_1 1$ for all $h \in M_1$ and $h > 0$. For $h \in M_1$ and $h < 0$ this result follows from the relation $\overset{*}{\sin}(-h) = \overset{*}{\sin} h$. Thus for all $h \in M_1$ and $h \neq 0$ we have $\frac{\sin^* h}{h} =_1 1$. We shall see later that this is the non-standard form of the relation $\lim\limits_{x \to 0} \frac{\sin x}{x} = 1$.

# CHAPTER 2.

# THE THEORY OF LIMITS OF SEQUENCES

## 1. Convergent sequences.

Let $\{s_n : n \in N\}$ be a standard sequence, i.e., a mapping of $N$ into $R$. Its non-standard extension is a mapping of $N^*$ into $R^*$. This mapping will be denoted by $\{s^*_n : n \in N^*\}$. We shall often write, however, $s_n$ in place of $s^*_n$ whenever $n \in N$. From Theorem 7.7 of Chapter 1, it follows that $\{s_n : n \in N\}^* = \{s^*_n : n \in N^*\}$. Observe, however, that for some $\omega \in N^* - N$ $s^*_\omega = s$, where $s$ is a standard number. In fact, this occurs if and only if the sequence repeats the number $s$ an infinite number of times.

Recall that by definition a standard sequence $\{s_n : n \in N\}$ is convergent if there exists a standard number $s$ such that

$$(\forall \epsilon)(\epsilon \in R_+ \Rightarrow (\exists n_o)(n_o \in N \text{ and } (\forall n)(n \in N \text{ and } n \geq n_o \Rightarrow |s_n - s| < \epsilon))).$$

If the sequence $\{s_n : n \in N\}$ is convergent, then $s$ is called its limit and we write $s = \lim_{n \to \infty} s_n$. It follows from Chapter 1, Theorem 7.3 that removing, changing or adding a finite number of elements of a sequence $\{s_n : n \in N\}$ does not affect the set $\{s^*_\omega : \omega \in N^* - N\}$. Hence, if a sequence $\{s_n : n \in N\}$ is convergent with limit $s$, then $|s^*_\omega - s| < \epsilon$ for all $\epsilon \in R_+$ and all $\omega \in N^* - N$, i.e., $|s^*_\omega - s| \in M_1$ or equivalently $s^*_\omega =_1 s$ for all $\omega \in N^* - N$. This suggests the following theorem.

THEOREM 1.1. <u>A standard sequence</u> $\{s_n : n \in N\}$ <u>is convergent with limit</u> $s$, <u>if and only if</u> $s^*_\omega =_1 s$ <u>for all</u> $\omega \in N^* - N$.

PROOF. We have only to show that the condition is sufficient. To this end assume that $s^*_\omega =_1 s$ for all $\omega \in N^* - N$ and for some $s \in R$. If not $(\lim_{n\to\infty} s_n = s)$ holds, then there exist an injection $\Omega$ of $N$ into $N$ and a number $\mathcal{E}_o \in R_+$ such that $|s_{\Omega(n)} - s| > \mathcal{E}_o$ for all $n \in N$. Hence, if $\Omega \in \omega$, then $s^*_\omega \neq_1 s$. Since the range of $\Omega$ is an infinite subset of $N$ we have $\omega \in N^* - N$. This contradicts the assumption.

REMARKS 1. The condition $s^*_\omega =_1 s$ for all $\omega \in N^* - N$ is equivalent to $\mathrm{st}(s^*_\omega) = s$ for all $\omega \in N^* - N$. Hence, if a sequence is convergent, then its limit is uniquely determined.

2. Since any change which affects only a finite number of elements of a given sequence $\{s_n : n \in N\}$ does not affect the set $\{s^*_\omega : \omega \in N^* - N\}$ (Chapter 1, Theorem 7.3), Theorem 1.1 implies that if a sequence is convergent, then any other sequence which is obtained from this sequence by changing only a finite number of elements of the sequence is also convergent with the same limit.

3. Theorem 1.1 implies also immediately that every subsequence of a convergent sequence is convergent and converges to the same limit.

THEOREM 1.2 <u>Every convergent sequence of real numbers is bounded.</u>

PROOF. If $\lim_{n\to\infty} s_n = \varepsilon$ holds, then by Theorem 1.1, $s^*_\omega \in M_0$ for all $\omega \in N^* - N$. Hence, by Theorem 7.4 of Chapter 1, $\{s_n : n \in N\}$ is bounded.

EXAMPLES 1. We have $\lim_{n\to\infty} (1/n) = 0$. Indeed, for all $\omega \in N^* - N$ we have that $1/\omega =_1 0$. In the same way, one shows that $\lim_{n\to\infty} (1/n)^p = 0$ $(p > 0)$.

2. We have $\lim_{n\to\infty} a^n = 0$ if $0 < |a| < 1$. Indeed, since $(1/a)^\omega$ is infinitely large for all $\omega \in N^* - N$ we obtain that $a^\omega =_1 0$ for all $\omega \in N^* - N$.

3. We have $\lim_{n\to\infty} a^{(1/n)} = 1$ if $a > 1$. Indeed, let $s_n = a^{(1/n)} - 1$, then $a = (1 + s_n)^n$. Hence, $a > ns_n$ for all $n \in N$. We conclude that $a > \omega s^*_\omega$ for all $\omega \in N^* - N$ or equivalently $0 < s^*_\omega < (a/\omega) \in M_1$, i.e., $s^*_\omega =_1 0$ for all $\omega \in N^* - N$. If $0 < a < 1$, then observe that $(1/a)^{(1/\omega)} =_1 1$ or $1 =_1 a^{(1/\omega)}$ for all $\omega \in N^* - N$. If $a = 1$, the limit relation is trivial.

4. We have $\lim_{n\to\infty} n^{(1/n)} = 1$. Indeed, observe that $n = (1+s_n)^n$, where $s_n = \sqrt[n]{n} - 1$. From $n = (1+s_n)^n$ it follows that $2n \geq n(n-1)s_n^2$. Hence, $0 < s^*_\omega \leq (4/(\omega-1))^{1/2} \in M_1$, i.e., $s^*_\omega =_1 0$ for all $\omega \in N^* - N$.

REMARK. We have supplied these simple examples only to show the reader that the non-standard form for convergence replaces in an elegant way the $\varepsilon$ and index arguments.

THEOREM 1.3 <u>If a montone standard sequence</u> $\{s_n : n \in N\}$ <u>has the property that for some</u> $\omega \in N^* - N$, $s^*_\omega \in M_0$, <u>then the sequence is convergent with limit</u> $st(s^*_\omega)$.

PROOF. Assume that the sequence is increasing, i.e., $s_1 \leq s_2 \leq \cdots \leq s_n \leq \cdots$. Then its non-standard extension is also increasing. Now assume that for some $\omega \in N^* - N$, $s^*_\omega \in M_0$. Then $s_n \leq st(s^*_\omega)$ for all $n \in N$. Hence, $st(s^*_{\omega'}) \leq st(s^*_\omega)$ for <u>all</u> $\omega' \in N^* - N$. But if $st(s^*_{\omega'}) < st(s^*_\omega)$, then the same argument shows that $st(s^*_\omega) \leq st(s^*_{\omega'})$. We conclude that $st(s^*_{\omega'}) = st(s^*_\omega)$ for all $\omega' \in N^* - N$, i.e., $s^*_{\omega'} =_1 st(s^*_\omega)$ for all $\omega' \in N^* - N$, which completes the proof of the theorem.

REMARK. One recognizes immediately in Theorem 1.3 the well-known theorem that bounded monotone sequences of real numbers are convergent.

## 2. The Algebra of Limits

In this section we shall give non-standard proofs of the usual rules for calculating with limits.

THEOREM 2.1 <u>Let</u> $\{u_n : n \in N\}$ <u>and</u> $\{v_n : n \in N\}$ <u>be two standard sequences. Assume that</u> $\lim_{n\to\infty} u_n = u$ <u>and</u> $\lim_{n\to\infty} v_n = v$.

<u>Then we have also</u>

(i) $\lim_{n\to\infty} s_n = u + v$ <u>and</u> $\lim_{n\to\infty} d_n = u-v$, <u>where</u> $s_n = u_n + v_n$ <u>and</u> $d_n = u_n - v_n$ <u>for all</u> $n \in N$.

(ii) $\lim_{n\to\infty} p_n = uv$, <u>where</u> $p_n = u_n v_n$ <u>for all</u> $n \in N$,

(iii) $\lim_{n\to\infty} q_n = u/v$ <u>if</u> $v \neq 0$, <u>where</u> $q_n = u_n/v_n$ <u>for all</u> $n \in N$.

PROOF. (i) By Theorem 1.1 we have $u^*_\omega =_1 u$ and $v^*_\omega =_1 v$ for all $\omega \in N^* - N$. Since $s^*_\omega = u^*_\omega + v^*_\omega$ and $d^*_\omega = u^*_\omega - v^*_\omega$, the result follows immediately

(ii) Observe that $p^*_\omega = u^*_\omega v^*_\omega =_1 uv$ for all $\omega \in N^* - N$.

(iii) Observe that $q^*_\omega = u^*_\omega/v^*_\omega =_1 u/v$ for all $\omega \in N^* - N$, if $v \neq 0$.

3. <u>Cauchy's Criterion for Convergence</u>.

We shall discuss in this section the form which Cauchy's criterion takes on in non-standard analysis.

Recall that a standard sequences $\{s_n : n \in N\}$ is called a <u>Cauchy sequence</u> if it satisfies the following condition:

$$(\forall \varepsilon)\ (\varepsilon \in R_+ \Rightarrow (\exists\ n)\ (n \in N \text{ and } (\forall\ p)(\forall\ q)(p \geq n,\ q \geq n \Rightarrow |s_p - s_q| < \varepsilon))).$$

If we then use the fact, expressed in Theorem 7.3 of Chapter 1, that the set $\{s^*_\omega : \omega \in N^* - N\}$ for a sequence $\{s_n : n \in N\}$ is unchanged if we remove or replace a finite number of elements of that sequence we obtain that $|s^*_\omega - s^*_{\omega'}| < \varepsilon$ for all $\varepsilon > 0$ and $\omega, \omega' \in N^* - N$. Hence, $s^*_\omega =_1 s^*_\omega$ for all $\omega, \omega' \in N^* - N$; and this is apparently the form to which the condition of being a Cauchy sequence reduces in non-standard analysis. Indeed, we shall prove the following

THEOREM 3.1. <u>A standard sequence</u> $\{s_n : n \in N\}$ <u>converges if and only if</u> $s^*_\omega =_1 s^*_{\omega'}$ <u>for all</u> $\omega, \omega' \in N^* - N$.

PROOF. If the sequence $\{s_n : n \in N\}$ converges, then there exists a standard number $s$ such that $s^*_\omega =_1 s$ for all $\omega \in N^* - N$ (Theorem 1.1). Hence, $s^*_\omega =_1 s^*_{\omega'}$ for all $\omega, \omega' \in N^* - N$. Conversely, assume that $s^*_\omega =_1 s^*_{\omega'}$ for all $\omega, \omega' \in N^* - N$. If $s^*_\omega$ is finite for some $\omega \in N^*$, then $s^*_{\omega'} =_1 s$ for all $\omega' \in N^* - N$, where $s = st(s^*_\omega)$; and hence, by Theorem 1.1, the sequence converges. Assume therefore that $s^*_\omega$ is infinitely large for all $\omega \in N^* - N$. Then the sequence $\{s_n : n \in N\}$ is unbounded (Chapter 1, Theorem 7.6). Hence, for every $\omega \in N^* - N$ there exists an element $\omega' \in N^* - N$ such that $|s^*_\omega + 1| < |s^*_{\omega'}|$. This contradicts the assumption and finishes the proof of the theorem.

REMARKS 1. Incidentally, this theorem shows that conversely, the condition $s^*_\omega =_1 s^*_{\omega'}$ for all $\omega, \omega' \in N^* - N$ implies that the sequence is a Cauchy sequence.

2. The condition which expresses that a standard sequence is a Cauchy sequence has been given by different authors in different ways. They are all easily checked to be equivalent. It may be of interest to remark that Theorem 3.1 may also be used to prove such equivalences. We shall illustrate this by means of the following example. E. Goursat in his now classical "Cours d'Analyse Mathématique" writes that a (standard) sequence $\{s_n : n \in N\}$ is a Cauchy sequence if it satisfies the following condition:

$$(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists n)(n \in N \text{ and } (\forall m)(m \in N \Rightarrow |s_{n+m} - s_n| < \varepsilon))).$$

One sees immediately that this statement implies the following statement.

$(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists \omega_o)(\omega_o \in N^* - N \text{ and } (\forall \omega)(\omega \in N^* - N \Rightarrow |s^*_\omega - s^*_{\omega_o}| < \varepsilon)))$.

Hence, $s^*_\omega - s^*_{\omega_o} =_1 0$ for all $\omega \in N^* - N$, i.e., $s^*_\omega =_1 s^*_{\omega'}$ for all $\omega, \omega' \in N^* - N$. Thus Goursat's condition reduces immediately to the non-standard form of the statement that a sequence is a Cauchy sequence.

## 4. The Existence of Generalized Limits in the Sense of S. Banach.

Let $\ell_\infty(N)$ as usual denote the linear space of all bounded real sequences $\{s_n : n \in N\}$. A real function on $\ell_\infty(N)$ will be called a Hahn-Banach limit and will be denoted by $\underset{n\to\infty}{\mathrm{Lim}}\, s_n$ if it satisfies the following conditions:

(i) $\underset{n\to\infty}{\mathrm{Lim}}(as_n+bt_n) = a\,\underset{n\to\infty}{\mathrm{Lim}}\, s_n + b\,\underset{n\to\infty}{\mathrm{Lim}}\, t_n$,

(ii) $\underset{n\to\infty}{\mathrm{Lim}}\, s_n \geq 0$ whenever $s_n \geq 0$ for all $n \in N$,

(iii) $\underset{n\to\infty}{\lim}\inf s_n \leq \underset{n\to\infty}{\mathrm{Lim}}\, s_n \leq \underset{n\to\infty}{\lim}\sup s_n$.

A Hahn-Banach limit is called a Banach-Mazur limit if in addition to the conditions (i), (ii) and (iii) it satisfies also the following condition:

(iv) $\underset{n\to\infty}{\mathrm{Lim}}\, s_{n+k} = \underset{n\to\infty}{\mathrm{Lim}}\, s_n$ for all $k \in N$

Thus a Banach-Mazur limit is a Hahn-Banach limit which is invariant under the shift operation.

Observe that, in particular, condition (iii) implies that if $\underset{n\to\infty}{\lim}\, s_n$ exists in the usual sense, then $\underset{n\to\infty}{\mathrm{Lim}}\, s_n = \underset{n\to\infty}{\lim}\, s_n$.

We shall now first prove with the help of non-standard analysis that there do exist Hahn-Banach limits.

THEOREM 4.1. For every $\omega \in N^* - N$, the function $\underset{n\to\infty}{\mathrm{Lim}}\, s_n = st(s^*_\omega)$ which is defined for all elements of $\ell_\infty$ is a Hahn-Banach limit.

PROOF. Since for every $\omega \in N^* - N$ and every bounded real sequence $\{s_n : n \in N\}$, $s^*_\omega$ is finite it follows that the function mentioned in the theorem is well-defined. We shall now prove that it satisfies the conditions (i) - (iii).

(i) Observe that if $u_n = as_n + bt_n$ $(n\in N)$, then $u^*_\omega = as^*_\omega + bt^*_\omega$ for all $\omega \in N^* - N$ and hence, $\underset{n\to\infty}{\mathrm{Lim}}\,(as_n+bt_n) = a\,\underset{n\to\infty}{\mathrm{Lim}}\,s_n + b\,\underset{n\to\infty}{\mathrm{Lim}}\,t_n$.

(ii) If $s_n \geq 0$ for all $n \in N$, then $s^*_\omega \geq 0$ for all $\omega \in N^* - N$. Hence, $st(s^*_\omega) \geq 0$ for all $\omega \in N^* - N$, which proves (ii).

(iii) If we set $p_n = \inf\,(s_{n+m} : m\in N)$, $n \in N$, and $q_n = \sup\,(s_{n+m} : m\in N)$, $(n\in N)$, then $p_n \leq s_n \leq q_n$ for all $n \in N$. Hence, for all $\omega \in N^* - N$ we have $p^*_\omega \leq s^*_\omega \leq q^*_\omega$. We conclude that $st(p^*_\omega) \leq st(s^*_\omega) \leq st(q^*_\omega)$. Since $st(p^*_\omega) = \underset{n\to\infty}{\lim}\inf\, s_n$ and $st(q^*_\omega) = \underset{n\to\infty}{\lim}\sup\, s_n$, the proof is completed.

REMARKS. 1. Observe that if a Hahn-Banach limit $\underset{N\to\infty}{\mathrm{Lim}}\,s_n$ on $\ell_\infty$ is of the form $\underset{n\to\infty}{\mathrm{Lim}}\,s_n = st(s^*_\omega)$ for some $\omega \in N^* - N$, then it has also the following property: $\underset{n\to\infty}{\mathrm{Lim}}\,s_n t_n = (\underset{n\to\infty}{\mathrm{Lim}}\,s_n)(\underset{n\to\infty}{\mathrm{Lim}}\,t_n)$, i.e., such a Hahn-Banach limit is also multiplicative.

2. It is not without interest to observe that for every bounded real sequence $\{s_n : n\in N\}$ there exist $\omega, \omega' \in N^* - N$ such that $\underset{n\to\infty}{\lim}\inf\, s_n = st(s^*_\omega)$ and $\underset{n\to\infty}{\lim}\sup\, s_n = st(s^*_{\omega'})$.

THEOREM 4.2. <u>For every</u> $\omega \in N^* - N$, <u>the function</u> $\underset{n\to\infty}{\mathrm{Lim}}\,s_n = st\left(\sum_{k=o}^{\omega} s_k/\omega\right)$ <u>which is defined for all elements of</u> $\ell_\infty$ <u>is a Banach-Mazur limit</u>.

PROOF. For every bounded real sequence $\{s_n : n\in N\}$ the sequence $a_n = \dfrac{s_o+\cdots+s_n}{n+1}$, $n \in N$, is bounded and hence for all $\omega \in N^* - N$,

$a^*_\omega = \sum_{k=o}^{\omega} s_k/\omega$ is finite. This shows that the expression given in the theorem defines a function on $\ell_\infty$. It is easy to verify that

$\operatorname*{Lim}_{n\to\infty} s_n = st\left(\sum_{k=o}^{\omega} s_k/\omega\right)$ is a Hahn-Banach limit. In order to prove that it is a Banach-Mazur limit we observe that

$$\sum_{k=o}^{\omega} s_{k+n}/\omega = \left(\sum_{k=o}^{\omega} s_k\right)/\omega + \left(\sum_{k=\omega+1}^{\omega+n} s_k\right)/\omega - \left(\sum_{k=o}^{n-1} s_k\right)/\omega =_1 \left(\sum_{k=o}^{\omega} s_k\right)/\omega,$$ and the required result follows.

REMARK. In contrast to the properties of a Hahn-Banach limit a non-zero Banach-Mazur limit is <u>not</u> multiplicative. Indeed, let $s_{2n-1} = 0$ $(n=1,2,\ldots)$ and $s_{2n} = 1$ $(n\in N)$ and let $s'_n = s_{n+1}$ $(n\in N)$. Then $s_n s'_n = 0$ for all $n \in N$ and $s_n + s'_n = 1$ for all $n \in N$. If $\operatorname*{Lim}_{n\to\infty}$ is a Banach-Mazur limit such that $\operatorname*{Lim}_{n\to\infty} 1 = 1$, then $\operatorname*{Lim}_{n\to\infty}$ is multiplicative will imply that $0 = \operatorname{Lim}(s_n s'_n) = (\operatorname*{Lim}_{n\to\infty} s_n)(\operatorname*{Lim}_{n\to\infty} s'_n)$ and $\operatorname*{Lim}_{n\to\infty} s_n + \operatorname*{Lim}_{n\to\infty} s'_n = 1$. But this contradicts the fact that $\operatorname*{Lim}_{n\to\infty} s_n = \operatorname*{Lim}_{n\to\infty} s'_n$.

Incidentally, this shows that not every Hahn-Banach limit is multiplicative.

## 5. <u>A Theorem of G. Pólya and G. Szegö.</u>

The following problem occurs in the famous book "Aufgaben und Lehrsätze aus der Analysis I" by G. Pólya and G. Szegö (problem 99, p.17 and solution p. 171)

<u>Let</u> $\{s_n : n\in N\}$ <u>be a real sequence such that</u> $|s_{n+m} - s_n - s_m| \le s$ <u>for all</u> $n,m \in N$. <u>Show that</u> $\lim_{n\to\infty} s_n/n = \sigma$ <u>exists and that</u> $|s_n - \sigma n| \le s$ <u>for all</u> $n \in N$.

The purpose of this section is to give a new non-standard proof of a slightly more general statement. The proposed proof is considerably simpler than the one given by Pólya and Szegö in their book.

THEOREM 5.1. Let $\{s_n : n \in N\}$ be a standard sequence such that $|s_{n+m}-s_n-s_m| \leq s(n^p + m^p)$ for all $n, m \in N$ and for some $0 \leq p < 1$. Then $\lim_{n\to\infty} \frac{s_n}{n} = \sigma$ exists and $|s_n - \sigma n| \leq s(\frac{n^p}{1-2^{p-1}})$.

PROOF. We shall first show that under the hypothesis of the theorem the following relation holds:

$$(*) \quad \left| \frac{s_{2^k n}}{2^k} - s_n \right| \leq sn^p \sum_{i=0}^{k-1} 2^{i(p-1)} \text{ for all } n \in N \text{ and } k \in N.$$

It is obvious that (*) holds for $k = 1$. Assume that (*) holds for k.

Then $\left|\frac{s_{2^k 2n}}{2^k} - s_{2n}\right| \leq s^p 2n^p \sum_{i=0}^{k-1} 2^{i(p-1)} = 2sn^p \sum_{i=1}^{k} 2^{i(p-1)}$, and hence,

we have $$\left|\frac{s_{2^{k+1}n}}{2^{k+1}} - s_n\right| \leq \left|\frac{s_{2^{k+1}n}}{2^{k+1}} - \frac{s_{2n}}{2}\right| + \left|\frac{s_{2n}}{2} - s_n\right| \leq s \cdot n^p \sum_{i=0}^{k} 2^{i(p-1)}.$$

We conclude from the principle of induction that (*) holds. Hence, if

$\omega \in N^* - N$, then $$\left|\frac{s^*_{2^\omega n}}{2^\omega} - s_n\right| \leq sn^p \sum_{i=0}^{\omega} 2^{i(p-1)} = sn^p \frac{1-2^{(p-1)\omega}}{1-2^{p-1}}.$$ Since

$p < 1$ implies $2^{(p-1)\omega} \in M_1$, we have $\frac{s^*_{2^\omega n}}{2^\omega} \in M_o$ for all $n \in N$. Observe now that the hypothesis of the theorem implies that

$$\left| s^*_{2^\omega(m+n)} - s^*_{2^\omega m} - s^*_{2^\omega n} \right| \leq s\ 2^{\omega p}\ (n^p+m^p).$$ Hence, we obtain

$$\left| \frac{s^*_{2^\omega(n+m)}}{2^\omega} - \frac{s^*_{2^\omega m}}{2^\omega} - \frac{s^*_{2^\omega n}}{2^\omega} \right| \leq s\ 2^{\omega(p-1)}(n^p+m^p) \in M_1.$$ Since $\frac{s^*_{2^\omega n}}{2^\omega} \in M_o$

for all $n \in N$, we obtain, by setting $t_n = st(\frac{s^*_{2^\omega n}}{2^\omega})$ and by taking the standard part of each side of the last inequality, that $t_{n+m} = t_n + t_m$, i.e. $t_n = nt_1 = n\sigma$, where $\sigma = st(\frac{s^*_{2^\omega}}{2^\omega})$. If we now take the standard part of of each side of the inequality $\left| \frac{s^*_{2^\omega n}}{2^\omega} - s_n \right| \leq s\ n^p \frac{1-2^{(p-1)\omega}}{1-2^{p-1}}$ we obtain $|n\sigma - s_n| \leq \frac{sn^p}{1-2^{p-1}}$. This shows that $\lim_{n\to\infty} \frac{s_n}{n}$ exists and is equal to $\sigma$.

This completes the proof of the theorem.

REMARKS 1. If instead of the condition $|s_{n+m}-s_n-s_m| \leq s(n^p + m^p)$ for all $n, m \in N$ we assume $|s_{n+m}-s_n-s_m| \leq s\ (|n|^p + |m|^p)$ for all integers $n, m$, then we have that $\lim_{n\to\pm\infty} \frac{s_n}{n} = \sigma$ and $|s_n-\sigma n| \leq \frac{s|n|^p}{1-2^{p-1}}$. The proof remains the same. Incidentally, for $p = o$, the latter form of the theorem was the one given by Pólya and Szegö.

2. The constant $\sigma$ is also unique in the following sense: If for some constant $\sigma'$, $|s_n-n\sigma'| \leq s'n^{p'}$, where $0 \leq p' < 1$, then $\sigma = \sigma'$. Indeed, in this case $|\sigma n - \sigma' n'| \leq s''(n^p+n^{p'})$, $n \in N$.

## 6. The Theory of Limits of Double Sequences

In this section we shall briefly discuss the theory of limits of double sequences from a non-standard viewpoint.

Recall that a double sequence is a mapping of $N \times N$ into $R$. Usually, double sequences are denoted by $\{s_{m,n} : (m,n) \in N \times N\}$, where $s_{m,n}$ are the elements of the double sequences. Instead of denoting the elements of the double sequence by $s_{m,n}$ we shall denote them by $s(m,n)$ which is more convenient for our considerations.

Since so far we have only introduced the non-standard extension of a function of <u>one</u> real variable we shall give a separate definition for the non-standard extension of a double sequence. This definition is contained in a more general definition to be given later in Chapter 4.

DEFINITION 6.1 (<u>Non-standard extension of a double sequence</u>). <u>Let</u> $\{s(m,n) : (m,n) \in N \times N\}$ <u>be a double sequence of real numbers. Then the double sequence</u> $\{s^*(a,b) : (a,b) \in N^* \times N^*\}$ <u>of non-standard numbers defined as follows:</u>

$$S \in s^*(a,b) \Leftrightarrow (\exists A)(\exists B)(A \in a \text{ and } B \in b \text{ and } \{n : S(n) = s(A(n), B(n))\} \in \mathcal{U})$$

<u>is called the non-standard extension of</u> $\{s(m,n) : (m,n) \in N \times N\}$.

It is evident that $(\forall S)(\exists A)(\exists B)$ $(S \in s^*(a,b)$ and $A \in a$ and $B \in b$ and $\{n : S(n) = s(A(n), B(n))\} \in \mathcal{U}) \Rightarrow (\forall S)(\forall A)(\forall B)(S \in s^*(a,b)$ and $A \in a$ and $B \in b$ and $\{n : S(n) = s(A(n), B(n))\} \in \mathcal{U})$. Furthermore,

$\{s(m,n): (m,n) \in N \times N\}^* = \{s^*(a,b):(a,b) \in N^* \times N^*\}$.

Recall that a double sequence $\{s(m,n):(m,n) \in N \times N\}$ is <u>convergent</u> if and only if there exists a number $s \in R$ such that

$$(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists n_o)(n_o \in N \text{ and } (\forall_n)(\forall_m)(m,n \in N,\ m,n \geq n_o \Rightarrow$$

$$|s(m,n) - s| < \varepsilon))).$$

In that case, $s$ is called the <u>limit</u> of the double sequence and is denoted by $\lim_{m,n\to\infty} s(m,n)$.

Thus, if a double sequence $\{s(m,n): (m,n) \in N \times N\}$ is convergent with limit $s$, then $(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\forall \omega)(\forall \omega')(\omega,\omega \in N^* - N \Rightarrow$ $|s^*(\omega,\omega') - s| < \varepsilon))$. Hence, $s^*(\omega,\omega') =_1 s$ for all $\omega, \omega' \in N^* - N$. The converse holds also. Indeed, assume that $s^*(\omega,\omega') =_1 s$ for all $\omega,\omega' \in N^* - N$ and that not $(\lim_{m,n\to\infty} s(m,n) = s)$ hold. Then there exist injections $\Omega, \Omega'$ of $N$ into $N'$ and a number $\varepsilon_o > o$ such that $|s(\Omega(n), \Omega'(n)) - s| > \varepsilon_o$ for all $n \in N$. Hence, $s^*(\omega, \omega') \neq_1 s$, where $\Omega \in \omega$ and $\Omega' \in \omega'$, which contradicts our assumption. We have therefore analogously to Theorem 1.1.

THEOREM 6.1. <u>A double sequence</u> $\{s(m,n):(m,n) \in N \times N\}$ <u>is convergent with limit</u> $s$ <u>if and only if</u> $s^*(\omega,\omega') =_1 s$ <u>for all</u> $\omega, \omega' \in N^* - N$.

From Theorem 7.4, Chapter 1 we deduce then immediately the following theorem.

THEOREM 6.2. <u>Every convergent double sequence is bounded.</u>

Cauchy's criterion for convergence takes on the following form.

THEOREM 6.3. <u>A double sequence</u> $\{s(m,n):(m,n) \in N \times N\}$ <u>is convergent if and only if</u> $s^*(\omega, \omega') =_1 s^*(\mu,\mu')$ <u>for all</u> $\omega, \omega', \mu,\mu' \in N^* - N$.

PROOF. If a double sequence is convergent, then the condition of Theorem 6.3 follows immediately from the condition of Theorem 6.1. To prove the converse, argue in exactly the same way as in the corresponding part of the proof of Theorem 3.1.

An important role in the theory of limits of double sequences play the so-called <u>repeated limits</u> $\lim_{n\to\infty} (\lim_{m\to\infty} s(m,n))$ and $\lim_{m\to\infty} (\lim_{n\to\infty} s(m,n))$. If a double sequence is convergent, then its repeated limits may <u>not</u> exist. In this case, however, we can prove that as much as the following result holds.

THEOREM 6.4. <u>If the double sequence</u> $\{s(m,n):(m,n) \in N \times N$ <u>is convergent, then for every sequence</u> $\{\omega_m:m \in N\}$ <u>of elements of</u> $N^* - N$ <u>we have</u> $\lim_{m\to\infty} st(s^*(m,\omega_m))$ <u>exists and is equal to the limit of the double sequence. The same result holds for the sequence</u> $st(s^*(\omega_n,n))$.

PROOF. If the double sequence $\{s(m,n):(m,n) \in N \times N\}$ is convergent with limit s, then $|s(mn) - s| < \varepsilon$ for all sufficiently large m, n. Hence, $|s^*(m,\omega_m)-s| \leq \varepsilon$ for sufficiently large m. If we take the standard part of each side of the last in equality, then we obtain $|st(s^*(m,\omega_m))-s| < \varepsilon$ for sufficiently large m. We conclude that $\lim_{m\to\infty} st(s^*(m,\omega_m)) = s$. This completes the proof of the theorem.

REMARK. Theorem 6.4 implies in particular that $\lim_{n\to\infty} (\liminf_{m\to\infty} s(m,n))$ $= \lim_{n\to\infty} (\limsup_{m\to\infty} s(m,n)) = \lim_{m\to\infty} (\liminf_{n\to\infty} s(m,n)) = \lim_{m\to\infty} (\limsup_{n\to\infty} s(m,n)) = s$, provided the double sequence $\{s(m,n):(m,n) \in N \times N\}$ is convergent with limit $s$.

As an immediate corollary to Theorem 6.4 we have the following theorem.

THEOREM 6.5. <u>If the double sequence</u> $\{s(m,n):(m,n) \in N \times N$ <u>is convergent with limit</u> $s$, <u>then</u> $\lim_{n\to\infty} (\lim_{m\to\infty} s(m,n))$ <u>exists and is equal to</u> $s$ <u>if and only if</u> $\lim_{m\to\infty} s(m,n)$ <u>exists for all</u> $n \in N$.

PROOF. Observe that $\lim_{m\to\infty} s(m,n) = st(s^*(\omega,n))$ for all $n \in N$ and all $\omega \in N^* - N$. Apply then Theorem 6.4 in order to obtain the required result.

In the following theorem the following condition is of importance. Let $\{s(m,n):(m,n) \in N \times N\}$ be a double sequence. We say that $\lim_{m\to\infty} s(m,n)$ exists <u>uniformly</u> in $n$ if the following condition is satisfied:

$$(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists m_0)(m_0 \in N \text{ and } (\forall m)(\forall n)(N \ni m,m' \geq m_0,\ n \in N$$
$$\Rightarrow |s(m,n) - s(m',n)| < \varepsilon ))).$$

Hence, for every $n \in N^*$, $s^*(\omega,n) =_{.} s^*(\omega',n)$ for all $\omega, \omega' \in N^* - N$. The converse holds also. Indeed, assume that $(\forall n)(\forall \omega)(\forall \omega')(n \in N, \omega,\omega' \in N^*-N$ $\Rightarrow s^*(\omega,n) =_1 s^*(\omega'n))$ and not $(\lim_{m\to\infty} s(m,n)$ uniformly in $n)$. Then the latter condition implies that there exist injections $\Omega$, $\Omega'$ and $\Omega''$ of $N$ into $N$ and a number $\varepsilon_o \in R_+$ such that $|s(\Omega'(n), \Omega(n)) - s(\Omega''(n),\Omega(n))| \geq \varepsilon_o > o$ for all $n \in N$. Hence $s^*(\omega',\omega) \neq_1 s^*(\omega'',\omega)$ where $\Omega \in \omega$, $\Omega' \in \omega'$ and $\Omega'' \in \omega''$. This contradicts the assumption. We have proved therefore the following theorem.

THEOREM 6.6. <u>Let</u> $\{s(m,n):(m,n) \in N \times N\}$ <u>be a double sequence</u>. <u>Then</u> $\lim_{m\to\infty} s(m,n)$ <u>exists uniformly in</u> $n$ <u>if and only if for every</u> $n \in N^*$, $s^*(\omega,n) =_1 s^*(\omega',n)$ <u>for all</u> $\omega, \omega' \in N^* - N$.

We shall use this theorem in the proof of the following theorem.

THEOREM 6.7 <u>Let</u> $\{s(m,n):(m,n) \in N \times N\}$ <u>be a double sequence</u>. <u>Then we have</u> $\lim_{m,n\to\infty} s(m,n) = \lim_{m\to\infty} (\lim_{n\to\infty} s(m,n)) = \lim_{n\to\infty} (\lim_{m\to\infty} s(m,n))$ <u>if and only if the following two conditions are satisfied</u>:

(i) $\lim_{m\to\infty} s(m,n)$ <u>exists uniformly in</u> $n$,

(ii) $\lim_{n\to\infty} s(m,n)$ <u>exists for every</u> $m \in N$.

PROOF. The conditions are necessary. Indeed, (ii) follows immediately from the existence of the repeated limits. In order to prove (i), observe that $\lim_{n\to\infty} (\lim_{m\to\infty} s(m,n))$ exists implies that for every $n \in N$, $s^*(\omega,n) =_1 s^*(\omega', n)$ for all $\omega,\omega' \in N^* - N$. From the existence of the double limit it follows in particular, using Theorem 6.3, that for every $\omega_o \in N^* - N$, $s^*(\omega,\omega_o) =_1 s^*(\omega',\omega_o)$ for all $\omega,\omega' \in N^* - N$. If we combine these two results together we obtain that for every $n \in N^*$, $s^*(\omega,n) =_1 s^*(\omega',n)$ for all $\omega, \omega' \in N^* - N$, i.e., (i) by Theorem 6.7.

We shall turn now to the proof that the conditions are sufficient. We put $\lim\limits_{m\to\infty} s(m,n) = s(n)$. Then, by Theorem 6.7 and condition (i), we have $s^*(\omega,\omega') - s^*(\omega') =_1 0$ for all $\omega, \omega' \in N^* - N$. Hence, by Theorem 6.1, $\lim\limits_{m,n\to\infty} (s(m,n) - s(n)) = 0$. If we use Theorem 6.4, we obtain that for every $\omega \in N^* - N$, $\lim\limits_{m\to\infty} st(s^*(m,\omega)) = st(s^*(\omega))$. Since (ii) implies that for all $\omega,\omega' \in N^* - N$, $st(s^*(m,\omega)) = st(s^*(m,\omega'))$ we have $s^*(\omega) =_1 s^*(\omega')$ for all $\omega, \omega' \in N^* - N$. Then $s^*(\omega,\omega') =_1 s^*(\omega')$ and $s^*(\mu,\mu') =_1 s^*(\mu')$ implies $s^*(\omega,\omega') =_1 s^*(\mu,\mu')$ for all $\omega,\omega',\mu,\mu' \in N^* - N$, i.e., by Theorem 6.3, $\lim\limits_{m,n\to\infty} s(m,n)$ exists. In this case $\lim\limits_{m,n\to\infty} s(m,n) = \lim\limits_{n\to\infty} s(n) = \lim\limits_{n\to\infty} (\lim\limits_{m\to\infty} s(m,n))$. The remainder of the theorem follows from Theorem 6.5. This completes the proof of the theorem.

REMARK. Of course in the conditions of Theorem 6.7 we may interchange $n$ and $m$.

We shall conclude this section with a theorem due to U. Dini which gives a useful sufficient condition for the existence of the double limit provided the repeated limits exist and are equal.

THEOREM 6.8 (Dini). *Let* $\{s(m,n) : (m,n) \in N \times N\}$ *be a double sequence such that* $\lim\limits_{n\to\infty} (\lim\limits_{m\to\infty} s(m,n)) = \lim\limits_{m\to\infty} (\lim\limits_{n\to\infty} s(m,n)) = s$. *If for every* $m \in N$,

$s(m,n)$ _is increasing in_ $n$ _or for every_ $m \in N$, $s(m,n)$ _is decreasing in_ $n$, _then_ $\lim_{m,n \to \infty} s(m,n)$ _exists and is equal to_ $s$.

PROOF. There is no loss in generality if we assume that for every $m \in N$, $s(m,n)$ is decreasing in $n$ and $\lim_{m \to \infty} s(m,n) = 0$ for all $n \in N$. By Theorem 6.4, $\lim_{n \to \infty} (\lim_{m \to \infty} s(m,n)) = 0$ implies that $\lim_{n \to \infty} st(s^*(\omega,n)) = 0$ for all $\omega \in N^*-N$. Since for all $\omega, \omega' \in N^* - N$ and all $n \in N$ we have $0 \leq s^*(\omega,\omega') \leq s^*(\omega,n)$ we obtain that $st(s^*(\omega,\omega')) = 0$, for all $\omega,\omega' \in N^*-N$, i.e., $s^*(\omega,\omega') =_1 0$ for all $\omega,\omega' \in N^* - N$. Hence, by Theorem 6.1, $\lim_{m,n \to \infty} s(m,n) = 0$. This completes the proof of the theorem.

7. The Theory of Infinite Series.

Let $\{a_k : k \in N\}$ be a sequence of real numbers. The sequence $\{s_n : n \in N\}$, where $s_n = \sum_{k=0}^{n} a_k$ $(n \in N)$ is called an _infinite series_ and for this sequence the symbolic notation $\sum_{k=0}^{\infty} a_k$ is used. The elements of the sequences $\{s_n : n \in N\}$ are called the _partial sums_ of the infinite series. We have the following simple relations: $a_n = s_n - s_{n-1}$ for all $n \in N$ and $\sum_{k=p}^{q} a_k = s_q - s_{p-1}$ for all $p, q \in N$ with $p \leq q$.

An infinite series is said to be _convergent_ if the sequence of its partial sums is convergent. The limit of this sequence is called the _sum_ of the infinite series. In symbols: If $\sum_{k=0}^{\infty} a_k$ is an infinite series and

$\lim_{n\to\infty} s_n = s$, then we write $\sum_{k=0}^{\infty} a_k = s$.

From Theorem 1.1 it follows then immediately

THEOREM 7.1. An infinite series $\sum_{k=0}^{\infty} a_k$ is convergent if and only if there exists a standard number $s$ such that $s^*_\omega =_1 s$ for all $\omega \in N^* - N$.

If an infinite series $\sum_{k=0}^{\infty} a_k$ is not convergent, then for every $\omega \in N^* - N$, $s^*_\omega$ is still well defined, and will be denoted symbolically by $\sum_{k=0}^{\omega} a_k$.

Cauchy's criterion takes on the following form

THEOREM 7.2. An infinite series $\sum_{k=0}^{\infty} a_k$ is convergent if and only if $s^*_\omega =_1 s^*_{\omega'}$ for all $\omega, \omega' \in N^* - N$.

From the relation $a_n = s_n - s_{n-1}$, $n \in N$ we deduce now immediately the following theorem

THEOREM 7.3. If an infinite series $\sum_{k=0}^{\infty} a_k$ is convergent, then $\lim_{n\to\infty} a_n = 0$.

PROOF. From $a_n = s_n - s_{n-1}$ we deduce that $a^*_\omega = s^*_\omega - s^*_{\omega-1}$. If the infinite series is convergent, then by Theorem 7.2, $s^*_\omega - s^*_{\omega-1} =_1 0$ for all $\omega \in N^* - N$. Thus the required result follows from Theorem 1.1.

EXAMPLES 1. Show that $\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1$. For every $\omega \in N^* - N$, we have $\sum_{k=1}^{\omega} \frac{1}{k(k+1)} = \sum_{k=1}^{\omega} \frac{1}{k} - \sum_{k=1}^{\omega} \frac{1}{k+1} = \sum_{k=1}^{\omega} \frac{1}{k} - \sum_{k=2}^{\omega+1} \frac{1}{k} = 1 - \frac{1}{\omega+1} =_1 1$.

2. (The geometric series) Let $x$ be any real number. The infinite series $\sum_{k=o}^{\infty} x^k$ is the so-called geometric series. If $x \neq 1$, then its partial sums can be expressed as follows: $s_n = \frac{1-x^{n+1}}{1-x}$ $(n \in N)$. Hence, $s^*_\omega = \frac{1-x^{\omega+1}}{1-x} = \frac{1}{1-x} - \frac{x^{\omega+1}}{1-x}$ if $x \neq 1$. We conclude that for all $|x| < 1$, $s^*_\omega =_1 \frac{1}{1-x}$ for all $\omega \in N^* - N$.

3. If we replace $x$ by $-x$ in the geometric series, then we obtain the infinite series $\sum_{k=o}^{\infty} (-1)^k x^k$. Hence, for all $\omega \in N^* - N$ we have $\sum_{k=o}^{\omega} (-1)^k x^k = \frac{1-(-1)^{\omega+1} x^{\omega+1}}{1+x} = \frac{1}{1+x} - \frac{(-1)^{\omega+1} x^{\omega+1}}{1+x}$ provided $x \neq -1$. If we put $x = 1$, then we obtain $\sum_{k=o}^{\omega} (-1)^k = \frac{1}{2} + \frac{(-1)^\omega}{2}$. We conclude, $st(\sum_{k=o}^{\omega} (-1)^k) = 1$ if $\omega$ is even, and $st(\sum_{k=o}^{\omega} (-1)^k) = 0$ if $\omega$ is odd. If we substitute $x = 1 - \frac{1}{\omega+1}$, we obtain $\sum_{k=o}^{\omega} (-1)^k (1 - \frac{1}{\omega+1})^k = \frac{1+(-1)^\omega (1-\frac{1}{\omega+1})^{\omega+1}}{2 - \frac{1}{\omega+1}}$

Since $(1 - \frac{1}{\omega+1})^{\omega+1} =_1 \frac{1}{e}$ we obtain that for all $\omega \in N^* - N$, $st(\sum_{k=o}^{\omega} (-1)^k (1-\frac{1}{\omega+1})^k) = \frac{1}{2}$.

The latter statement expresses that the infinite series $\sum_{k=o}^{\infty} (-1)^k$ is Abel summable with Abel sum $\frac{1}{2}$.

THEOREM 7.3. *Let* $\sum_{k=0}^{\infty} a_k$ *be an infinite series such that* $a_k \geq 0$ *for all* $k \in N$. *Then* $\sum_{k=0}^{\infty} a_k$ *is convergent if and only if for some* $\omega \in N^* - N$, $\sum_{k=0}^{\omega} a_k$ *is finite.*

PROOF. If $a_k \geq 0$ for all $k \in N$, then the sequence $\{s_n : n \in N\}$ is increasing. Hence, the required result follows from Theorem 1.3.

The following theorem is a useful complement to the preceeding theorem.

THEOREM 7.4. *Let* $\sum_{k=0}^{\infty} a_k$ *be an infinite series such that* $a_k \geq 0$ *for all* $k \in N$. *Then* $\sum_{k=0}^{\infty} a_k$ *is divergent if and only if for some* $\omega \in N^* - N$, $\sum_{k=0}^{\omega} a_k$ *is infinitely large.*

EXAMPLES 1. (*Harmonic series*) The harmonic series $\sum_{k=1}^{\infty} \frac{1}{k}$ is divergent. Indeed, $s_{2^{n+1}} - s_{2^n} = \sum_{k=2^n}^{2^{n+1}} \frac{1}{k} = \sum_{k=1}^{2^n} \frac{1}{2^n+k} > \frac{2^n}{2^{n+1}} = \frac{1}{2}$. Hence, $s^*_{2^\omega} > \frac{\omega}{2}$.

2. The infinite series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ is convergent if $p > 1$ and divergent if $p \leq 1$. Observe that if $p > 1$, $s_{2^{n+1}} - s_{2^n} = \sum_{k=1}^{2^n} \frac{1}{(k+2^n)^p} < 2^{(1-p)n}$. Hence, $s^*_{2^\omega} < \frac{1}{1-2^{1-p}}$. If $p \leq 1$, then $s_{2^{n+1}} - s_{2^n} = \sum_{k=1}^{2^n} \frac{1}{(k+2^n)^p}$

$$> \frac{2^n}{2^{(n+1)p}} > \frac{1}{2^p} .$$

Hence, $s^*_{2^\omega} > \frac{\omega}{2^p}$ .

3. The infinite series $\sum_{k=2}^{\infty} \frac{1}{k \log k}$ is divergent. Indeed,

$s_{2^{n+1}} - s_{2^n} = \sum_{k=1}^{2^n} \frac{1}{(k+2^n)\log(k+2^n)} > \frac{1}{(2 \log 2)(n+1)}$ . Hence

$s^*_{2^\omega} > \frac{1}{2 \log 2} \sum_{k=1}^{2^\omega} \frac{1}{k+1}$ .

From Theorems 7.3 and 7.4 we deduce immediately the following theorem.

THEOREM 7.5 If $a_k \geq 0$, $b_k \geq 0$, $a_k \leq b_k$, $k \in N$, then $\sum_{k=0}^{\infty} a_k$ converges if $\sum_{k=0}^{\infty} b_k$ converges and $\sum_{k=0}^{\infty} b_k$ diverges if $\sum_{k=0}^{\infty} a_k$ diverges.

As another example of non-standard methods we shall prove the following well-known theorem.

THEOREM 7.6 (Olivier) If $a_k \geq 0$ and $a_k \geq a_{k+1}$, $k \in N$, then $\sum_{k=0}^{\infty} a_k$ converges implies that $\lim_{k \to \infty} k\, a_k = 0$.

PROOF. Let $\omega \in N^* - N$ and let $\omega' = [\frac{\omega}{2}]$, i.e., $\omega'$ is the largest infinitely large natural number $< \frac{\omega}{2}$ . Then $s^*_\omega - s^*_{\omega'} \geq \frac{\omega}{2} a^*_\omega \geq 0$. Hence, $\omega\, a^*_\omega =_1 0$ for all $\omega \in N^* - N$, which completes the proof of the theorem.

An infinite series $\sum_{k=0}^{\infty} a_k$ is said to be absolutely convergent of the infinite series $\sum_{k=0}^{\infty} |a_k|$ is convergent. If we denote the partial sums of the series $\sum_{k=o}^{\infty} |a_k|$ by $|s|_n$, $n \in N$, then we see immediately that $|s^*_\omega - s^*_{\omega'}| \leq ||s|^*_\omega - |s|^*_{\omega'}|$ . Hence, we have by Theorem 7.2.

THEOREM 7.7. An absolutely convergent infinite series is convergent.

An important tool in the theory of convergence for series with arbitrary terms in Abel's summation by parts.

If $\{a_k : k \in N\}$ and $\{b_k : k \in N\}$ are two sequences of real numbers, then

$$\sum_{k=p}^{q} a_k b_k = \sum_{k=p}^{q} (s_k - s_{k-1})b_k = \sum_{k=p}^{q} s_k b_k - \sum_{k=p}^{q} s_{k-1} b_k =$$

$$= \sum_{k=p}^{q} s_k b_k - \sum_{k=p-1}^{q-1} s_k b_{k+1} = \sum_{k=p}^{q} s_k(b_k - b_{k+1}) - s_{p-1}b_p + s_q b_{q+1},$$

where $s_{-1}$ is to be understood to be equal to zero.

Thus, if $p = o$ and $q = \omega$, where $\omega \in N^* - N$, we obtain

$$\sum_{k=o}^{\omega} a_k b_k = \sum_{k=o}^{\omega} s_k(b_k - b_{k+1}) + s^*_\omega b^*_{\omega+1} .$$

Various theorems can be deduced from this relation. As an example we shall prove the following theorem.

THEOREM 7.8.(du Bois-Reymond) If the infinite series $\sum_{k=o}^{\infty} (b_k - b_{k+1})$ is absolutely convergent and if the infinite series $\sum_{k=o}^{\infty} a_k$ is convergent, then the infinite series $\sum_{k=o}^{\infty} a_k b_k$ is convergent.

PROOF. Let $\omega < \omega'$, $\omega, \omega' \in N^* - N$, then by Abel's summation by parts we have $\sum_{k=\omega}^{\omega'} a_k b_k = \sum_{k=\omega}^{\omega'} s_k(b_k - b_{k+1}) - s^*_{\omega-1} b^*_{\omega} + s^*_{\omega'-1} b^*_{\omega'+1}$. Now $\sum_{k=0}^{\infty} (b_k - b_{k+1})$ converges absolutely implies that $b^*_{\omega} = b^*_{\omega'}$ for all $\omega, \omega' \in N^* - N$. Furthermore, the sequence $\{s_k : k \in N\}$ is bounded and $s^*_{\omega} = s^*_{\omega'}$ for all $\omega, \omega' \in N^* - N$. Thus we obtain immediately that $\sum_{k=\omega}^{\omega'} a_k b_k =_1 0$ for all $\omega, \omega' \in N^* - N$.

Let $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ be two infinite series. Then for all $\omega \in N^* - N$, we have

$$\left(\sum_{k=0}^{\omega} a_k\right) \left(\sum_{k=0}^{\omega} b_k\right) = \sum_{k=0}^{\omega} \left(\sum_{j=0}^{k} a_j b_{k-j}\right) + \sum_{k=0}^{\omega} \left(\sum_{j=0}^{k} a^*_{\omega-k+j} b^*_{\omega-j}\right).$$

The first part of the right hand side of this equality is nothing but the usual Cauchy product of two infinite series. The second term may be considered as a correction term. It is well-known that this correction term may fail to be infinitesimal for all $\omega \in N^* - N$ even if the series $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ are convergent. It is immediate that $\left(\sum_{k=0}^{\infty} a_k\right) \left(\sum_{k=0}^{\infty} b_k\right) = \sum_{j=0}^{\infty} \left(\sum_{j=0}^{\infty} a_j b_{k-j}\right)$ if and only if for all $\omega \in N^* - N$ $\sum_{k=0}^{\omega} \left(\sum_{j=0}^{k} a^*_{\omega-k+j} b^*_{\omega-j}\right) =_1 0$, provided that the series $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ are convergent.

THEOREM 7.9. (<u>Mertens</u>) <u>If</u> $\sum_{k=0}^{\infty} a_k$ <u>converges absolutely and</u> $\sum_{k=0}^{\infty} b_k$ <u>is convergent, then</u> $\left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{k=0}^{\infty} b_k\right) = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{\infty} a_j b_{k-j}\right)$.

PROOF. Define $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$, $C_n = \sum_{k=0}^{n}\left(\sum_{j=0}^{k} a_j\, b_{k-j}\right)$, $\lim_{n\to\infty} A_n = A$ and $\lim_{n\to\infty} B_n = B$. Then $C_n = A_n B - \sum_{k=0}^{n} a_k(B-B_{n-k})$ for all $n \in N$. To complete the proof we have to show that $\sum_{k=0}^{\omega} a_k^* \,(B-B^*_{\omega-k}) =_1 0$ for all $\omega \in N^* - N$. Observe that $\left|\sum_{k=0}^{\omega} a_k^* \,(B-B^*_{\omega-k})\right| \leq \sum_{k=0}^{\omega} |a_k^*| \cdot M$. Hence, by taking standard parts and observing that for all $0 \leq k \leq \omega$, $st(a_k^*(B-B^*_{\omega-k})) = 0$ we obtain the required result.

# CHAPTER 3

# THE THEORY OF REAL FUNCTIONS

## 1. The Topology of R in Non-Standard Analysis.

In this section we shall discuss the main properties of the interval topology in R. To this end, we shall recall the following important definition of point set topology.

(i) (Adherent point) Let S be a subset of R a point $s \in R$ is called an adherent point of S if every neighborhood of s has a point in common with S.

(ii) (Closure of a set) A set $S \subseteq R$ together with all its adherent points is called the closure of S in R and is denoted by $\bar{S}$.

(iii) (Isolated point) A point $s \in R$ is called an isolated point of $S \subseteq R$ if $s \in S$ and if there exists a neighborhood of s which has no points in common with S other than s.

An isolated point of a set is an adherent point of a set. An adherent point of a set which is not an isolated point of a set is called a non-trivial adherent point. Isolated points are often referred to as trivial adherent points.

(iv) (Interior point) A point $s \in R$ is called an interior point of

$S$ *if there exists a neighborhood of* $s$ *which is entirely contained in* $S$.

Every interior point of a set is a non-trivial adherent point of that set.

In the theorems below we shall give the non-standard forms of the above definitions.

THEOREM 1.1. *Let* $S \subseteq R$, *then* $s \in \bar{S}$ *if and only if there exists an element* $a \in S^*$ *such that* $s = st(a)$.

PROOF. Assume that $s \in \bar{S}$. Then, $(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists s')(s' \in S$ and $|s'-s| < \varepsilon))$. Hence, $(\forall \varepsilon)(\varepsilon \in R_+^* \Rightarrow (\exists a)(a \in S^*$ and $|a-s| < \varepsilon))$. If we take for $\varepsilon$ an infinitesimal in the last statement, then we obtain that there exists an element $a \in S^*$ such that $|a-s| < \varepsilon$, i.e., $s = st(a)$.

Conversely, assume that $a \in S^*$ and that $a$ is finite, then we shall prove that $st(a) \in \bar{S}$. Given $\varepsilon > 0$, observe that $\{n\colon A(n) \in S\} \cap \{n\colon |A(n)-st(a)| < \varepsilon\} \in \mathcal{U}$, where $A \in a$. Hence, for every $\varepsilon > 0$ there exists an index $n \in N$ such that $A(n) \in S$ and $|A(n)-st(a)| < \varepsilon$, i.e., $st(a) \in \bar{S}$.

DEFINITION 1.1 (*Standard Part of a Set*) *Let* $E$ *be a subset of* $R^*$. *The set of all standard parts of the finite elements of* $E$ *is called the standard part of* $E$ *and is denoted by* $st(E)$.

With this definition, Theorem 1.1 may be formulated as follows:

THEOREM 1.2. *Let* $S$ *be a subset of* $R$. *Then* $\bar{S} = st(S^*)$.

THEOREM 1.3. <u>A point</u> $s \in R$ <u>is a non-trivial adherent point of</u> $S \subseteq R$ <u>if and only if there exists an element</u> $a \in S^* - S$ <u>such that</u> $a =_1 s$, <u>or equivalently, there exists an infinitesimal</u> $h \in M_1$ <u>such that</u> $h \neq 0$ <u>and</u> $s + h \in S^*$.

PROOF. Let $s$ be a non-trivial adherent point of $S$. Then $(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists s')(s' \in S$ and $s' \neq s$ and $|s - s'| < \varepsilon))$. Hence, if we take for $\varepsilon$ an infinitesimal $> 0$, then there exists an element $a \in S^*$ such that $a \neq s$ and $a =_1 s$. Then $a \in S$ contradicts $a =_1 s$, i.e., $a \in S^* - S$. Conversely, assume that $a$ is a finite element of $S^* - S$. We have to show that $s = st(a)$ is a non-trivial adherent point of $S$. The statement $s$ is a non-trivial adherent point of $S$ is equivalent to the statement $s \in \overline{S - \{s\}}$. Now, by Theorem 1.2 and Theorem 7.3 of Chap. 1, $\overline{S - \{s\}} = st((S - \{s\})^*) = st(S^* - \{s\})$. Since $a \in S^* - S$ and the only standard elements of $S^*$ are in $S$ (Theorem 7.1 (v), Chap. 1), we obtain that $a \in S^* - \{s\}$. Hence, $st(a) \in \overline{S - \{s\}}$

As an immediate corollary we have the following theorem.

THEOREM 1.4. <u>A point</u> $s \in R$ <u>is an isolated point of</u> $S \subseteq R$ <u>if and only if</u> $s \in S^*$ <u>and</u> $(\forall a)(a \in M_0$ <u>and</u> $a \in S^* - S \Rightarrow s \neq st(a))$ <u>or equivalently</u> $s + h \notin S^*$ <u>for all</u> $0 \neq h \in M_1$.

The set of all non-trivial adherent points of a set $S \subseteq R$ is called its <u>derived set</u> and is denoted by $S'$.

THEOREM 1.5. <u>If</u> $S \subseteq R$, <u>then</u> $S' = st(S^* - S)$. <u>The set of all its isolated points is the set</u> $S - st(S^* - S)$.

Interior points or inner points of a subset of R can be characterized in a non-standard way as follows:

THEOREM 1.6. <u>A point</u> $s \in R$ <u>is an interior point of</u> $S \subseteq R$ <u>if and only if for every</u> $h \in M_1$, $s + h \in S^*$.

PROOF. Let $S$ be a subset of $R$ and let $s \in R$ such that for some $h \in M_1$, $s + h \notin S^*$. Then $s + h \in C_{R^*}(S^*)$, where $C_{R^*}(S^*)$ means the complement of $S^*$ in $R^*$. But $C_{R^*}(S^*) = (C_R(S))^*$ (Theorem 7.1 (iv), Chap. 1). Hence, $s$ is a non-trivial adherent point of $C_R(S)$, i.e., $s$ is not an interior point of $S$. Conversely, assume that $s + h \in S^*$ for all $h \in M_1$. Then for all $h \in M_1$, $s + h \notin (C_R(S))^*$. Thus $s$ is not an adherent point of $C_R(S)$, i.e., $s$ is an interior point of $S$.

We shall now prove the Bolzano-Weierstrass theorem.

THEOREM 1.7 (<u>Bolzano-Weierstrass</u>). <u>If</u> $S$ <u>is a bounded infinite subset of</u> $R$, <u>then</u> $S$ <u>has at least one non-trivial adherent point</u>.

PROOF. By Theorem 7.4 of Chap. 1, $S$ is infinite implies $S^*-S \neq \emptyset$. By Theorem 7.5 of Chap. 1, $S$ is bounded implies $S^* \subseteq M_0$. Let $a \in S^* - S$. Then $a \in M_0$ and hence, by Theorem 1.3, $s = st(a)$ is a non-trivial adherent point of $S$.

We shall give another proof of this theorem. Since $S$ is bounded, there exists a real number $u > 0$ such that $x \in S$ implies $|x| < u$. Since $S$ is infinite, we have that for every $n \in N$ there exists an interval of length $(2u)/n$ which contains infinitely many elements of $S$. We conclude that this holds in $N^*$ as well. Thus, given $\omega \in N^* - N$, there exists an interval of infinitesimal length $2u/\omega$ which contains infinitely many elements of $S^*$. Let $s$ be the standard part of the endpoints of that interval. Then $s$ is a non-trivial adherent point of $S$. Indeed, for infinitely many infinitesimals $h \neq 0$ we have that $s + h \in S^*$. Hence, by Theorem 1.3, $s$ is a non-trivial adherent point of $S$. This completes the proof of the theorem.

REMARK. It may be of interest to point out that conversely the Bolzano-Weierstrass theorem implies that $M_o/M_1$ is isomorphic to $R$ (Theorem 5.2 of Chapter 1). Indeed, the Bolzano-Weierstrass theorem is equivalent to the statement that bounded and closed subsets of $R$ are compact. Hence, if $a \in M_o$, then for every $A \in a$, $\lim_{\mathcal{U}} A$ exists and its value is the same for all elements $A \in a$. In order to prove this we have to observe that (i) the image of $\mathcal{U}$ under $A$ is a basis of an ultrafilter in $R$; (ii) since $A$ is bounded on a set of $\mathcal{U}$ and closed bounded subsets of $R$ are compact, the image of $\mathcal{U}$ under $A$ converges to an element of $R$ (a topological space is compact if and only if every ultrafilter is convergent). Furthermore, this limit is unique since $R$ is a Hausdorff space in its interval topology; (iii) if $A$ and $A'$ are elements of $a$, then they coincide on an element of $\mathcal{U}$. Hence $\lim_{\mathcal{U}} A = \lim_{\mathcal{U}} A'$.

If we set $h(a) = \lim_{\mathcal{U}} A$ for all $a \in M_o$, where $A \in a$, then it is easy to see that $h(a) = o$ if and only if $a \in M_1$ and that $h(a)$ has the same properties as the standard part homomorphism of $M_o$ onto $R$ with kernel $M_1$. Thus $h(a) = st(a)$ for all $a \in M_0$ or equivalently $M_0/M_1$ is isomorphic to $R$.

We have given an algebraic proof of the fact that $M_0/M_1$ and $R$ are isomorphic rather than the simple topological proof outlined above in order to make it possible to give non-standard proofs of the fundamental properties of the system of real numbers. Thus the proof given above for the Bolzano-Weierstrass theorem may be looked upon as a new proof of this fundamental theorem in analysis.

Needless to say that this proof involves the axiom of choice or at least the hypothesis that every proper filter is contained in an ultrafilter whereas the classical proof of the Bolzano-Weierstrass theorem does not make use of the set theoretic principles mentioned above.

## 2. Limits of Functions

Let $S \neq \emptyset$ be a subset of $R$ and let $f$ be a mapping of $S$ into $R$. If $s \in R$ is an adherent point of $S$, then we say that $f(x)$ converges to $\ell$ if $x$ tends to $s$ through $S$ if and only if $(\forall \varepsilon)\ (\varepsilon \in R_+ \Rightarrow (\exists \delta)(\delta > 0$ and $(\forall x)(x \in s$ and $0 < |x-s| < \delta \Rightarrow |f(x) - \ell| < \varepsilon))$. In that case we write $\lim_{S \ni x \to s} f(x) = \ell$. The reader should observe that if $s$ is an isolated point of $S$, then $\lim_{S \ni x \to s} f(x) = \ell$ holds for all $\ell \in R$. The interesting case is of course when $s$ is a non-trivial adherent point of $S$ and $f$ is not defined at $s$.

From this definition it follows immediately that for all infinitesimals $h \neq 0$ such that $s + h \in S^*$, $|f^*(s+h)-\ell| < \varepsilon$ for all $\varepsilon > 0$. Hence, $f^*(s+h) =_1 \ell$ for all infinitesimals $h \neq 0$ such that $s + h \in S^*$. This suggests that the following non-standard form of the previous definition holds.

THEOREM 2.1. <u>Let</u> $S$ <u>be a non-empty subset of</u> $R$ <u>and let</u> $f$ <u>be a mapping of</u> $S$ <u>into</u> $R$. <u>If</u> $s$ <u>is an adherent point of</u> $S$, <u>then</u> $\lim_{S \ni x \to s} f(x) = \ell$ <u>if and only if</u> $f^*(s+h) =_1 \ell$ <u>for all</u> $0 \neq h \in M_1$ <u>such that</u> $s + h \in S^*$.

PROOF. We have only to show that the given condition is sufficient. To this end, assume that non $(\lim_{S \ni x \to s} f(x) = \ell)$ holds. Then there exist an injection $\{s_n : n \in N\}$ of $N$ into $S$ and a number $\varepsilon \in R_+$ such that $s^*_\omega =_1 s$ for all $\omega \in N^* - N$ and $|f(s_n)-\ell| > \varepsilon_0$ for all $n \in N$. Hence $|f^*(s^*_\omega) - \ell| > \varepsilon_0$ for all $\omega \in N^* - N$. Since, $s^*_\omega = s + h$, where $0 \neq h \in M_1$ and $s + h \in S^*$ we obtain a contradiction.

REMARKS 1. If $\lim_{S \ni x \to s} f(x) = \ell$ exists, then $\ell$ is <u>uniquely</u> determined, provided $s$ is a <u>non-trivial adherent</u> point of $S$.

2. If $T$ is a subset of $S$ such that $s$ is an adherent point of $T$, then $\lim_{S \ni x \to s} f(x) = \ell$ implies $\lim_{T \ni x \to s} f(x) = \ell$. Indeed, by Chap. 1, Theorem 7.1 (iii), we have that $T^* \subset S^*$.

3. If $\lim_{S \ni x \to s} f(x) = \ell$, then there exists a neighborhood $V$ of $s$ such that $f$ is bounded on $V \cap S$. Indeed, there exists a neighborhood $V$ such that $f^*((V \cap S)^*) \subseteq M_0$.

4. Theorem 2.1 implies immediately the following well-known theorem: $\lim_{S \ni x \to s} f(x)$ exists if and only if for every sequence $\{s_n : n \in N\}$ such that $s_n \in S$ and $s_n \neq s$ $(n \in N)$ and $s = \lim_{n \to \infty} s_n$ we have $\lim_{n \to \infty} f(s_n)$ exists. The reader is advised to check this.

5. At the end of Chapter 1 we showed that $\frac{\sin^* h}{h} =_1 1$ for all $h \in M_1$. It follows now from Theorem 2.1 that this is equivalent to the well-known limit relation $\lim_{x \to 0} \frac{\sin x}{x} = 1$.

In the next theorem we shall give the usual rules for calculating with limits.

THEOREM 2.2. <u>Let</u> $f$ <u>and</u> $g$ <u>be two real functions defined on a non-empty subset</u> $S$ <u>of</u> $R$. <u>Let</u> $s \in R$ <u>be an adherent point of</u> $S$ <u>and assume that</u> $\lim_{S \ni x \to s} f(x) = \ell$ <u>and</u> $\lim_{S \ni x \to s} g(x) = m$. <u>Then we have also</u>

(i) $\lim_{S \ni x \to s} (f(x) \pm g(x)) = \ell \pm m$,

(ii) $\lim_{S \ni x \to s} f(x)g(x) = \ell m$,

(iii) $\lim_{S \ni x \to s} \frac{f(x)}{g(x)} = \frac{\ell}{m}$, <u>if</u> $m \neq 0$.

PROOF. (i) Observe that $(f \pm g)^* = f^* \pm g^*$. Hence, if $0 \neq h \in M_1$ such that $s + h \in S^*$, then $(f^* \pm g^*)(s+h) = f^*(s+h) \pm g^*(s+h) =_1 \ell \pm m$.

(ii) As in (i), $(fg)^* = f^* g^*$. Hence, if $0 \neq h \in M_1$ such that $s + h \in S^*$, then $(fg)^*(s+h) = f^*(s+h)\, g^*(s+h) =_1 \ell\, m$.

(iii) Observe, that $m \neq 0$ implies $g^*(s+h) \neq 0$ for all $0 \neq h \in M_1$ such that $s + h + S^*$. Hence, $(\frac{f}{g})^*(s+h) = \frac{f^*(s+h)}{g^*(s+h)} =_1 \frac{\ell}{m}$.

If the following theorem we shall give Cauchy's criterion for functions.

THEOREM 2.3. (Cauchy's criterion). Let $f$ be a real function defined on a non-empty subset $S$ of $R$. Let $s \in R$ be an adherent point of $S$. Then $\lim_{S \ni x \to s} f(x)$ exists as a finite number if and only if $f^*(s+h) =_1 f^*(s+k)$ for all $h, k \in M_1$ such that $s + h \in S^*$ and $s + k \in S^*$ and $h, k \neq 0$.

PROOF. That the condition is necessary follows immediately from Theorem 2.1. In order to prove that the condition is sufficient, observe that it implies that $f$ is bounded in some neighborhood of $s$. Then the required result follows immediately from Theorem 2.1 by taking $\ell = st(f^*(s+h))$.

We shall now discuss briefly in non-standard analysis the statement: $f(x)$ converges as $x$ tends to infinity. Let $f$ be a real function defined on a set $S \subseteq R$ which is not bounded above, i.e., for every $r \in R_+$ there exits an element $s \in S$ such that $s > r$. Then $\lim_{S \ni x \to \infty} f(x) = \ell$ if and only if $(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists r)(r \in R \text{ and } (\forall x)(x \in S \text{ and } x > r \Rightarrow |f(x) - \ell| < \varepsilon)))$.

Hence, $f^*(a) =_1 \ell$ for <u>all infinitely large positive numbers</u> $a \in S^*$. The converse holds also. In fact, we have

THEOREM 2.4. <u>Let</u> $f$ <u>be a real function defined on a non-empty subset</u> $S$ <u>of</u> $R$ <u>which is not bounded above</u>. <u>Then</u> $\lim_{S \ni x \to +\infty} f(x) = \ell$ <u>if and only if</u> $f^*(a) =_1 \ell$ <u>for all infinitely large positive numbers</u> $a \in S^*$.

The simple proofs of this theorem and the following theorem are left to the reader as an exercise.

Cauchy's criterion takes the following form.

THEOREM 2.5 (<u>Cauchy's criterion</u>) <u>Let</u> $f$ <u>be a real function defined on a non-empty subset</u> $S$ <u>of</u> $R$ <u>which is not bounded above</u>. <u>Then</u> $\lim_{S \ni x \to +\infty} f(x)$ <u>exists if and only if</u> $f^*(a) =_1 f^*(b)$ <u>for all infinitely large positive numbers</u> $a, b \in S^*$.

Similar theorems hold for the relation $\lim_{S \ni x \to -\infty} f(x) = \ell$ .

Let $S$ and $T$ be subsets of $R$. If $T$ is a set of adherent points of $S$, then $T^*$ is a set of adherent points of $S^*$ in $R^*$. Indeed, the latter statement is nothing but the non-standard form of the former statement. With this remark in mind we shall consider the following situation: Let $f$ be a real function defined on a subset $S \neq \emptyset$ of $R$. Let $T$ be a set of adherent points of $S$. Then $\lim_{S \ni x \to s} f(x)$ exists for all $s \in T$ if and only if

$$(\forall s)(s\in T \Rightarrow (\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists \delta)(\delta \in R_+ \text{ and}$$

$$(\forall x)(\forall y)(x\in S,\ y\in S,\ 0<|x-s|<\delta,\ 0<|y-s|<\delta \Rightarrow |f(x)-f(y)|<\varepsilon)))).$$

Hence in $R^*$ holds

$$(\forall s)(s\in T^* \Rightarrow (\forall \varepsilon)(\varepsilon \in R_+^* \Rightarrow (\exists \delta)(\delta \in R_+^* \text{ and}$$

$$(\forall x)(\forall y)(x\in S^*,\ y\in S^*,\ 0<|x-s|<\delta, 0<|y-s|<\delta \Rightarrow |f^*(x)-f^*(y)|<\varepsilon)))).$$

i.e., $\lim_{S^* \ni x \to s} f^*(x)$ exists for all $s \in T^*$. This proves the following theorem.

THEOREM 2.6. <u>Let</u> $f$ <u>be a real function defined on a subset</u> $S \neq \emptyset$ <u>of</u> $R$. <u>If</u> $T$ <u>is a set of adherent points of</u> $S$, <u>then</u> $\lim_{S \ni x \to s} f(x)$ <u>exists for all</u> $s \in T$ <u>implies</u> $\lim_{S^* \ni x \to s} f^*(x)$ <u>exists for all</u> $s \in T^*$.

3. <u>Continuity</u>.

Let $f$ be a real function defined on a subset $S$ of $R$. Recall that $f$ is said to be continuous at $s \in S$ if and only if $\lim_{S \ni x \to s} f(x) = f(s)$. Hence, by Theorem 2.1 we have

THEOREM 3.1. <u>Let</u> $f$ <u>be a real function defined on a non-empty subset</u> $S$ <u>of</u> $R$. <u>Then</u> $f$ <u>is continuous at</u> $s \in S$ <u>if and only if</u> $f^*(s+h) =_1 f(s)$ <u>for all</u> $h \in M_1$ <u>such that</u> $s + h \in S^*$, <u>or equivalently</u>, $f^*(a) =_1 f^*(b)$ <u>for all</u> $a,b \in S^*$ <u>such that</u> $st(a) = st(b) = s$.

As an immediate consequence of this theorem we have the following theorem.

THEOREM 3.2. *Let* $f$ *be a real function defined on a non-empty subset* $S$ *of* $R$. *Then* $f$ *is continuous on* $S$ *if and only if* $st(f^*(a)) = f(st(a))$ *for all* $a \in S^*$ *such that* $a$ *is finite*.

REMARK. Observe that the condition of the preceding theorem was shown already to hold for the elementary functions we treated as examples in section 7 of Chapter 1.

THEOREM 3.3. *Let* $f$ *and* $g$ *be two real functions defined on a non-empty subset* $S \subseteq R$ *and assume that* $f$ *and* $g$ *are continuous at* $s \in S$. *Then* $f + g$, $f - g$ *and* $f \cdot g$ *are continuous at* $s$. *The quotient* $f/g$ *is continuous at* $s \in S$ *provided that* $g(s) \neq 0$.

PROOF. Observe that $(f \pm g)^* = f^* \pm g^*$, $(fg)^* = f^*g^*$ and under the hypothesis of the theorem we have also $(f/g)^*(s+h) = (f^*/g^*)(s+h)$ for all $h \in M_1$.

For the composition of two real functions we have the following result.

THEOREM 3.4. *Let* $f$ *be a real function defined on a non-empty subset* $S$ *of* $R$ *and let* $g$ *be a real function defined on a subset* $T$ *of* $R$. *If* $f(S) \subset T$ *and* $f$ *is continuous at* $s \in S$ *and* $g$ *is continuous at* $f(s) \in T$, *then* $g \circ f = g(f)$ *is continuous at* $s$.

PROOF: Let $h \in M_1$ such that $s + h \in S^*$. Then $f^*(s+h) =_1 f(s)$. Since $g$ is continuous at $f(s)$ we have $g^*(f^*(s+h)) =_1 g(f(s))$. The required result follows then from the fact that $(g \circ f)^* = g^* \circ f^*$.

We shall give now a non-standard proof of the following important theorem which characterizes continuity.

THEOREM 3.5. Let $f$ be a real function defined on a non-empty subset $S$ of $R$. Then $f$ is continuous on $S$ if and only if $f^{-1}(T)$ is open in $S$ for every subset $T \subseteq f(s)$ which is open in $f(S)$.

PROOF. Assume first that $f$ is continuous on $S$ and that $T \subseteq f(S)$ is open in $f(S)$. In order to prove that $f^{-1}(T)$ is open in $S$ we have to show that for every $s \in f^{-1}(T)$, $s + h \in S^*$ for all $h \in M_1$ such that $s + h \in (f^{-1}(T))^*$. To this end, let $s \in f^{-1}(T)$ and $h \in M_1$ such that $s + h \in S^*$. Then, since $f$ is continuous at $s$ we have, by Theorem 3.1, that $f^*(s+h) =_1 f(s)$. But $T$ being open, the latter statement implies that $f^*(s + h) \in T^*$, i.e., $s + h \in (f^{-1}(T))^*$. To prove the converse, observe that by Theorem 7.7 of Chapter 1 we have $(f^{-1}(T))^* = (f^{-1})^*(T^*)$. Hence, under the hypothesis of the theorem, $f^*(s+h) \in T$, for all $h \in M_1$ such that $s + h \in S^*$. This holds for all subsets $T$ of $f(S)$ open in $f(S)$ and $f(s) \in T$, i.e., $f^*(s+h) =_1 f(s)$ for all $h \in M_1$ such that $s + h \in S^*$, which completes the proof of the theorem.

The following theorem follows immediately from Theorem 2.6.

THEOREM 3.6. Let $f$ be a real function defined on a subset $S \neq \emptyset$ of $R$. If $f$ is continuous on $S$, then $f^*$ is continuous on $S^*$ in the topology induced by the interval topology of $R^*$ on $S^*$.

We shall conclude this section with an application of the methods developed in this section.

THEOREM 3.7. *Let* $f$ *be a real function defined on a subset* $S \neq \emptyset$ *of* $R$. *If for every* $x \in S$, $\lim_{S \ni s \to x} f(s)$ *exists, then the function* $g$ *defined as follows:* $g(x) = f(x)$ *if* $x$ *is an isolated point of* $S$ *and* $g(x) = \lim_{S \ni s \to x} f(s)$ *if* $x$ *is a non-trivial adherent point of* $S$, *is continuous on* $S$.

PROOF. Let $x$ be a non-trivial adherent point of $S$. Then we have to show that $g^*(x+h) =_1 g(x)$ for all $h \in M_1$ such that $x + h \in S^*$. It follows from Theorem 2.6 that there exists an infinitesimal $k$ such that $x + h + k \in S^*$ and $f^*(x + h + k) =_1 g^*(x+h)$. From the hypothesis of the theorem it follows that $f^*(x+h+k) =_1 g(x)$. Hence, $g^*(x+h) =_1 g(x)$, i.e., $g$ is continuous at $x$.

## 4. Properties of Real Continuous Functions.

Let $f$ be a real function defined on a non-empty subset $S$ of $R$. Recall that $f$ is said to have an absolute maximum on the set $S$ if there exists a point $s \in S$ such that $f(x) \leq f(s)$ for all $x \in S$. If $s \in S$ and if there exists a neighborhood $V$ of $s$ such that $f(x) \leq f(s)$ for all $x \in S \cap V$, then $f$ is said to have a relative maximum at the point $s$. Absolute minimum and relative minimum are similarly defined, using $f(x) \geq f(s)$.

In non-standard analysis these notions can be formulated as follows:

THEOREM 4.1. *Let* $f$ *be a real function defined on a non-empty subset* $S$ *of* $R$.

(i) $f$ *has an absolute maximum on* $S$ *if and only if there exists a point* $s \in S$ *such that* $f^*(x) \leq f(s)$ *for all* $x \in S^*$

(ii) $f$ *has a relative maximum at* $s \in S$ *if and only if* $f^*(s+h) \leq f(s)$ *for all* $h \in M_1$ *such that* $s + h \in S^*$.

*For absolute minimum and relative minimum similar statements hold, using* $f(x) \geq f(s)$.

PROOF. We have only to show that the condition under (ii) is sufficient. To this end, assume that $f$ has not a relative maximum at $s \in S$. Then there exists a sequence $\{s_n : n \in N\}$ such that $s_n \in S$ for all $n \in N$, $\lim_{n \to \infty} s_n = s$ and $f(s_n) > f(s)$ for all $n \in N$. Hence, $f^*(s^*_\omega) > f(s)$ for all $\omega \in N^* - N$. Since $s^*_\omega \in S$ and $s^*_\omega =_1 s$ for all $\omega \in N^* - N$, we obtain a contradiction.

The following theorem is Weierstrass's famous theorem about the existence of an absolute maximum and an absolute minimum for a continuous function on a compact set.

THEOREM 4.2. (*Weierstrass*) *Let* $f$ *be a real function defined on a non-empty subset* $S$ *of* $R$. *If* $f$ *is continuous on* $S$ *and* $S$ *is bounded and closed, then* $f$ *is bounded and in fact, has an absolute maximum and an absolute minimum on* $S$.

PROOF. We shall first show that $f$ is bounded. For this purpose we have to show that $(f(S))^* \subseteq M_0$ (Chapter 1, Theorem 7.5). Let $a \in (f(S))^*$. Since $(f(S))^* = f^*(S^*)$ (Chapter 1, Theorem 7.7), there exists an element $b \in S^*$ such that $a = f^*(b)$. Hence, $S$ being bounded and closed, there exists an element $s \in S$ and an element $h \in M_1$ such that $b = s + h$. Since $f$ is continuous, we have $a = f^*(b) = f^*(s+h) =_1 f(s)$, i.e., $a$ is finite or equivalently $f$ is bounded.

In order to prove that $f$ attains its absolute maximum and its absolute minimum we have to show that $f(S)$ is closed. To this end let $\ell \in \overline{f(S)}$, i.e., there exists an element $b \in S^*$ such that $f^*(b) =_1 \ell$. Hence, $f(st(b)) = \ell$, i.e., $\ell \in f(S)$ which finishes the proof of the theorem.

## 5. The Intermediate Value Theorem for Continuous Functions.

The next two theorems deal with properties of real continuous functions defined on bounded and closed intervals in $R$. The first is Bolzano's famous theorem and the second is the intermediate value theorem which follows immediately from Bolzano's theorem.

The following proof of Bolzano's theorem is believed to be new.

THEOREM 5.1. (Bolzano) Let $f$ be a real continuous function defined on a bounded and closed interval $s \leq x \leq t$ $(s < t)$ of $R$. If $f(s)\, f(t) < 0$, then there exists at least one point $x$, $s < x < t$, such that $f(x) = 0$.

PROOF. Assume that $f(x) \neq 0$ for all $x$ such that $s < x < t$. Since $f(s)\, f(t) < 0$ implies that $f(s) \neq 0$ and $f(t) \neq 0$ we have, in fact, under this assumption that $f(x) \neq 0$ for all $x$ such that $s \leq x \leq t$. Under this hypothesis we shall prove that the following statement holds:

(*) *For every* $n \in N$ *such that* $n \geq 1$ *there exist points* $s_n, t_n$ *with the following properties:* $s \leq s_n < t_n \leq t$, $t_n - s_n = \frac{t-s}{n}$ *and* $f(t_n)/f(s_n) < 0$.

Indeed, consider the function $g(x) = f(x+(t-s)(n)/ f(x)$, $s \leq x \leq s + \frac{n-1}{n}(t-s)$. Then $\prod_{k=o}^{n-1} g(s+\frac{k}{n}(t-s)) = f(t)/f(s) < 0$, i.e., for some $k$, $0 \leq k \leq n-1$, we have $g(s+\frac{k}{n}(t-s)) < 0$, or equivalently, $f(s+\frac{k+1}{n}(t-s))/f(s+\frac{k}{n}(t-s)) < 0$. This completes the proof of (*).

In order to prove Bolzano's theorem, we deduce immediately from (*) that if $\omega \in N^* - N$, there exist a number $a \in R^*$ and a number $b \in R^*$ such that (i) $b-a = \frac{t-s}{\omega}$, (ii) $s \leq a < b \leq t$ and (iii) $f^*(b)/f^*(a) < 0$. From (i) and (ii) it follows that there exist infinitesimals $h$, $k$ and a number $x$, $s \leq x \leq t$ such that $a = x + h$ and $b = x + k$. Since $f$ is continuous, we have that $f^*(b) =_1 f(x)$ and $f^*(a) =_1 f(x)$. By hypothesis we have that $f(x) \neq 0$. Hence $st(f^*(b)/f^*(a)) = f(x)/f(x) = 1 \leq 0$, which is a contradiction and finishes the proof of the theorem.

The following theorem is an immediate consequence of Bolzano's theorem.

THEOREM 5.2 (*Intermediate value theorem*) *Let* $f$ *be a real continuous function defined on an interval* $I$ *of* $R$. *If* $s$ *and* $t$ *are two elements of* $I$ *such that* $s < t$ *and if* $c$ *is a real number which lies between*

$f(s)$ <u>and</u> $f(t)$, <u>then there exists a number</u> $u$ <u>with the following properties:</u> $s < u < t$ <u>and</u> $f(u) = c$.

PROOF. Apply Bolzano's theorem to the function $f(x) - c$ defined on the bounded and closed interval $s \leq x \leq t$.

6. <u>Uniform Continuity</u>.

Let $f$ be a real function defined on a non-empty set $S \subseteq R$. Recall that $f$ is said to be uniformly continuous on $S$ if the following statement holds: $(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists \delta)(\delta \in R_+$ and $(\forall x)(\forall y)(x \in S,\ y \in S$ and $|x-y| < \delta \Rightarrow |f(x)-f(y)| < \varepsilon)))$. Hence, if $a \in S^*$ and $b \in S^*$ such that $a =_1 b$, then $|f^*(a) - f^*(b)| < \varepsilon$ for all $\varepsilon > 0$, i.e., $f^*(a) =_1 f^*(b)$. This suggest the following theorem.

THEOREM 6.1 (<u>Uniform continuity</u>). <u>Let</u> $f$ <u>be a real function defined on a non-empty subset</u> $S$ <u>of</u> $R$. <u>Then</u> $f$ <u>is uniformly continuous on</u> $S$ <u>if and only if</u> $f^*(a) =_1 f^*(b)$ <u>for all</u> $a, b \in S^*$ <u>such that</u> $a =_1 b$.

PROOF. We have only to show that the condition is sufficient. If $f$ is <u>not</u> uniformly continuous, then there exist sequences $s_n$, $t_n$ $(n \in N)$ in $S$ and a positive number $\varepsilon_0$ such that $t_n - s_n$ tends to zero as $n$ tends to infinity and for all $n \in N$, $|f(t_n) - f(s_n)| > \varepsilon_0$. Hence, for all $\omega \in N^* - N$ we have $|f^*(t^*_\omega) - f^*(s^*_\omega)| > \varepsilon_0 > 0$ and $s^*_\omega =_1 t^*_\omega$. Since $s^*_\omega \in S^*$ and $t^*_\omega \in S^*$ we have obtained a contradiction.

We shall now give a simple non-standard proof of Heine's theorem.

THEOREM 6.2 (Heine) Let $f$ be a real function defined on a non-empty subset $S$ of $R$. If $S$ is bounded and closed, then $f$ is continuous on $S$ implies that $f$ is uniformly continuous on $S$.

PROOF. Assume that $a, b \in S^*$ and $a =_1 b$. Since $S$ is bounded and closed we have $s = st(a) = st(b) \in S$. The continuity of $f$ implies that $f^*(a) =_1 f(s)$ and $f^*(b) =_1 f(s)$, i.e., $f^*(a) =_1 f^*(b)$. Hence, by Theorem 6.1, $f$ is uniformly continuous on S. This completes the proof of the theorem.

REMARK. The reader is advised to compare the conditions given in Theorem 3.1 with the condition given in Theorem 6.1.

## 7. Additive Functions.

A real function $f$ defined on $R$ is said to be additive if it satisfies the following functional equation

$$f(x+y) = f(x) + f(y)$$

for all $x , y \in R$.

The function $f(x) = ax$, where $a \in R$, is an example of an additive function. It was shown by Cauchy that if $f$ is continuous and additive, then $f(x) = xf(1)$. The problem whether there exist non-continuous additive functions was open for a long time. It was solved in the affirmative by

Hamel in 1905 (see G. Hamel, Eine Basis aller Zahlen und die unstetige Lösungen der Funktional gleichung $f(x+y) = f(x) + f(y)$, Math. Ann. 60, 459-462 (1905)). In this paper Hamel gave the following solution. Consider R as a linear space over the field Q of the rational numbers. Then, by using the axiom of choice, he proved the existence of a basis, which is now called a Hamel basis, for this linear space R, i.e., he proved that there exists a non-empty set of real numbers H with the following two properties: (i) H is free, i.e., if $x_1, \dots, x_n$ are elements of H and $r_1, \dots, r_n$ are rationals, then $\sum_{i=1}^{n} r_i x_i = 0$ implies $r_i = 0$ for all $i = 1, 2, \dots, n$. (ii) For every real number $x \neq 0$ there exist elements $x_1, \dots, x_n$ in H and non-zero rationals $r_1, \dots, r_n$ such that $x = \sum_{i=1}^{n} r_i x_i$.

From (i) it follows immediately that the expansion of $x \neq 0$ in terms of elements of H is unique.

With a Hamel basis H for R we can now easily construct a non-continuous additive function. For this purpose let $a_o \in H$ be fixed. Then we set $f(x) = 0$ if $x = 0$ and if in the expansion $x = \sum_{i=1}^{n} r_i x_i$, the number $a_o$ does not occur. If in the expansion $\sum_{i=1}^{n} r_i x_i$ of $x \neq 0$, $a_o$ does occur with rational coefficient $r$, then we set $f(x) = r$. It is easy to see that f is additive. But $f(x) = xf(1)$ since $f \neq 0$ and $f(x) = 0$ for all $x \in H$ and $x \neq a_o$.

Since the appearance of Hamel's paper numerous articles on additive functions have appeared. They deal mainly with the problem to find conditions in order that an additive function satisfying these conditions is continuous. So it was shown that if an additive function is Lebesgue measurable, then it is of the form $ax$, where $a \in R$. The best result in this direction is due to Ostrowski. He obtained the following result: *If an additive function $f$ is bounded above on a Lebesgue measurable set of positive measure, then $f(x) = xf(1)$ for all $x \in R$.* It is easy to see that a real-valued Lebesgue measurable function has the above property.

Ostrowski's condition reduces to the condition that $f$ is bounded on some interval. In fact, it seems that most of the conditions given in order that an additive function is of the type $ax$, where $a \in R$, reduce to the latter condition. That is why we shall give now a non-standard proof of the following theorem, due to G. Darboux.

THEOREM 7.1. *Let $f$ be an additive function defined on $R$. If $f$ is bounded on some interval in $R$, then $f(x) = xf(1)$ for all $x \in R$.*

PROOF. If $f$ is additive on $R$, then its non-standard extension $f^*$ is additive on $R^*$. Furthermore, since the additivity of $f$ implies that $f(rx) = rf(x)$ for all $x \in R$ and all $r \in Q$, we have that $f^*(rx) = rf^*(x)$ for all $x \in R^*$ and $r \in Q^*$. Now assume that $f$ is bounded on some interval $I$ in $R$. Let $x_o$ be an interior point of $I$. Then there exists a constant $A$ such that $|f^*(x_o+h)| \leq A$ for all $h \in M_1$, i.e., $|f^*(h)| \leq A + |f(x_o)|$ for all $h \in M_1$. Since $nh \in M_1$ for all $h \in M_1$ and $n \in N$ we conclude that $|f^*(h)| \leq (A + |f(x_o)|)/n$ for all $n \in N$ and $h \in M_1$, i.e., $f^*(h) \in M_1$ for all $h \in M_1$. Thus, by Theorem 3.1, $f$ is continuous at $x_o$ and the result follows from Cauchy's theorem. Without resorting to Cauchy's theorem we can complete the proof as follows: Let $x \in R$. Then there exist a non-standard rational $r \in Q^*$ and an infinitesimal $h$ such that $x = r + h$. Hence $f(x) = f^*(r+h) = f^*(r) + f^*(h) = rf(1) + f^*(h)$, i.e., $f(x) = st(rf(1)) = xf(1)$. This completes the proof of the theorem.

REMARK. For further reading, the reader is referred to the following interesting paper: H. Kestelman, On the functional equation $f(x+y) = f(x)+f(y)$, Fund. Math., 34, 145-147 (1946).

## 8. Differentiation of Functions of a Real Variable.

Let $f$ be a real function defined on a non-empty subset $S$ of $R$. Recall that, if $s \in S$ is a non-trivial adherent point of $S$, then $f$ is *said to be differentiable at the point* $s \in S$ *if* $\lim\limits_{S \ni x \to s} \dfrac{f(x)-f(s)}{x-s}$ exists. Since $s$ is a non-trivial adherent point of $S$ this limit is uniquely determined. As is customary in mathematics this limit is denoted by $f'(s)$ and is called the *derivative of* $f$ *at* $s$.

The following theorem follows immediately from Theorem 2.1.

THEOREM 8.1. *Let* $f$ *be a real function defined on a non-empty subset* $S$ *of* $R$. *If* $s \in S$ *is a non-trivial adherent point of* $S$, *then* $f$ *is differentiable at* $s$ *if there exists a standard number, which is denoted by* $f'(s)$, *such that* $\dfrac{f^*(s+h) - f(s)}{h} =_1 f'(s)$ *for all* $h \in M_1$ *satisfying* $h \neq 0$ *and* $s + h \in S^*$.

If $f$ is differentiable at $s \in S$, then for all $h \in M_1$ such that $h \neq 0$ and $s + h \in S^*$ we have $f^*(s + h) - f(s) =_1 f'(s)h$. Since $f'(s)h \in M_1$ we obtain that $f^*(s+h) =_1 f(s)$ for all $h \in M_1$, $h \neq 0$ and $s + h \in S^*$, i.e., $f$ is continuous at $s$ (Theorem 3.1).

Thus we have shown

THEOREM 8.2. Let $f$ be a real function defined on a non-empty subset $S$ of $R$. If $s \in S$ is a non-trivial adherent point of $S$ and $f$ is differentiable at $s$, then $f$ is continuous at $s$.

REMARK. Let $f$ be differentiable at $x \in S$. If $dx$ denotes an infinitesimal such that $dx \neq 0$ and $x + dx \in S^*$, then, by the preceding theorem, $dy = f^*(x+dx) - f(x)$ is infinitesimal. The quotient $\frac{dy}{dx}$ is finite however and $f'(x) = st(\frac{dy}{dx})$. Thus we see that we are now able to give a precise meaning to Leibniz's original definition of differentiability.

From the algebra of limits the following theorem can easily be deduced.

THEOREM 8.3. If $f$ and $g$ are real functions defined on a non-empty set $S \subseteq R$, then $f + g$, $f - g$ and $f \cdot g$ are differentiable at those non-trivial adherent points of $S$ where $f$ and $g$ are differentiable. This holds also for $f/g$ provided $g \neq 0$ at those points. The derivatives are given by the following formulas: $(f \pm g)' = f' \pm g'$; $(fg)' = f'g + fg'$; $(f/g)' = (f'g - fg')/g^2$ $(g \neq 0)$.

The following theorem is an immediate consequence of Theorem 8.1.

THEOREM 8.4. Let $f$ be a real function defined on a non-empty subset $S$ of $R$. If $s \in S$ is a non-trivial adherent point of $S$, then $f$ is differentiable at $s$ if and only if there exists a standard number, which we denote by $f'(s)$, with the following property: for every $h \in M_1$, $h \neq 0$, and $s + h \in S^*$, there exists an infinitesimal, which we denote by $\varepsilon_f(s,h)$ or shortly $\varepsilon(s,h)$ if no confusion can arise, such that $f^*(s+h)-f(s) = hf'(s) + h\,\varepsilon(s,h)$.

PROOF. If $f^*(s+h) - f(s) = h\,f'(s) + h\,\varepsilon(s,h)$ holds for all $h \in M_1$, $h \neq 0$ and $s + h \in S^*$, where $\varepsilon \in M_1$, then $\frac{f^*(s+h)-f(s)}{h} = f'(s) + \varepsilon(s,h)$ $=_1 f(s)$. Conversely, if $f$ is differentiable at $s$, then the condition of the theorem holds with $\varepsilon(s,h) = \frac{f^*(s+h)-f(s)}{h} - f'(s)$.

## 9. The Chain Rule.

We shall give a simple proof now of Leibniz's rule for differentiating a composite function.

THEOREM 9.1 (Chain rule) Let $f$ be a real function defined on a non-empty subset $S$ of $R$. Assume that $s \in S$ is a non-trivial adherent point of $S$ and that $f$ is differentiable at $s$. Let $g$ be a real function defined on $f(S)$. If $f(s)$ is a non-trivial adherent point of $f(S)$ and $g$ is differentiable at $f(s)$, then the composite function $g(f) = g \circ f$ is differentiable at $s \in S$ and $(g(f(s)))' = g'(f(s))\,f'(s)$.

PROOF. Let $h \in M_1$, $h \neq 0$ and $s + h \in S^*$. Then $f^*(s+h) = f(s) + k$, and $k \in M_1$ since $f$ is continuous at $s$ (Theorem 8.2). Hence,

$g^*(f^*(s+h)) - g(f(s)) = g^*(f(s)+k) - g(f(s)) = k\, g'(f(s)) + k\, \varepsilon_g(f(s),k)$

$= (f^*(s+h) - f(s))\, g'(f(s)) + (f^*(s+h) - f(s))\, \varepsilon_g(f(s),k) = g'(f(s))f'(s)h +$
$h(g'(f(s))\varepsilon_f(s,h) + f'(s)\, \varepsilon_g(f(s),k) + k\, \varepsilon_f(s,h)\, \varepsilon_g(f(s),k)) = g'(f(s))f'(s)h +$
$h\eta$. It is easy to see now that $\eta \in M_1$. This completes the proof of the theorem.

10. Mean-Value Theorems.

In this section we shall discuss the following classical mean-value theorem of the differential calculus.

THEOREM 10.1 (Mean-value theorem). Let $f$ be a real function defined on a bounded and closed interval $I = I(s,t)$ of $R$, where $s$ and $t$ are its endpoints and $s < t$. Let $f$ be continuous on $I$ and differentiable for all $x$ such that $s < x < t$. Then there exists a standard number $u$ such that $s < u < t$ and $f(t) - f(s) = (t-s)\, f'(u)$.

This theorem is proved by means of Rolle's theorem. Rolle's theorem is that special case of the mean-value theorem which corresponds to the condition $f(t) = f(s)$. The general case is then deduced from this special case by considering the function $g(x) = f(x) - \frac{f(t) - f(s)}{t - s}\, x$.

Rolle's theorem is usually proved by applying the theorem of Weierstrass which states that a continuous function defined on a bounded and closed set attains its absolute maximum as well as its absolute minimum (Theorem 4.2). Now, if $f$ is continuous on the interval $I = I(s,t)$ and $f(s) = f(t)$, then

$f$ attains either its absolute maximum or its absolute minimum in the interior of $I$. Then, by using the fact that at an extreme point, the derivative of a differentiable function vanishes, the required result follows.

We propose to give another proof of Rolle's theorem which will not use the theorem of Weierstrass. We believe that this proof is new and more elementary. To this end, we first prove the following theorem which is due to P. Levy (1934).

THEOREM 10.2. (P. Levy). *Let* $f$ *be a real continuous function defined on a bounded and closed interval* $I = I(s,t)$, $s < t$. *If* $f(s) = f(t)$, *then for every* $n \geq 3$, *there exists an interval* $I_n = I(s_n, t_n)$, *such that* $s < s_n < t_n < t$, $t_n - s_n = \frac{t-s}{n}$ *and* $f(t_n) = f(s_n)$.

PROOF. Consider the continuous function $g(x) = f(x + \frac{t-s}{n}) - f(x)$, $s \leq x \leq s + \frac{n-1}{n}(t-s)$. Then $\sum_{k=0}^{n-1} g(s+\frac{k}{n}(t-s)) = f(t)-f(s) = 0$. Hence either $g(s+\frac{k}{n}(s-t)) = 0$ for all $k = 0, 1, 2, \ldots, n-1$, and then the proof is finished since $n \geq 3$, or $g$ takes on positive and negative values, say $g(s+\frac{p}{n}(s-t)) > 0$ and $g(s+\frac{q}{n}(t-s)) < 0$, where $0 \leq p < q \leq n-1$. Then, by the intermediate value theorem, $g(x) = 0$ for some $x$, $s + \frac{p}{n}(t-s) < x < s + \frac{q}{n}(t-s)$, which proves the theorem.

REMARK. For every $0 < c < 1$ and $c \neq \frac{1}{n}$ $(n = 1, 2, \ldots)$ there exists a continuous function $f(x)$ on the unit interval $0 \leq x \leq 1$ such that $f(0) = f(1)$ but $f(x+c)-f(x) \neq 0$ for all $0 \leq x \leq 1-c$.

Indeed, consider the function $f(x) = \sin^2 \frac{\pi x}{c} - x \sin^2 \frac{\pi}{c}$.

Levy's theorem is a consequence of the following more general statement: Let $f$ be a real continuous function defined on the bounded and closed interval $I = I(s,t)$ such that $f(s) = f(t)$. The set $C$ of all numbers $c$, such that $0 < c < t-s$ and $f(x+c) - f(x) \neq 0$ for all $x$ such that $s \leq x \leq t-c$ has the following property: If $c_1, c_2 \in C$, then $c_1 + c_2$, mod$(t-s)$, belongs to $C$.

We shall turn now to the <u>proof of Rolle's theorem</u>. From a repeated application of Levy's theorem it follows that there exist sequences $\{s_n : n \in N\}$, $\{t_n : n \in N\}$ of real numbers such that $s < s_n < s_{n+1} < t_{n+1} < t_n < t$, $t_{n+1} - s_{n+1} = \frac{t_n - s_n}{n}$ and $f(t_n) = f(s_n)$. Hence, there exists elements $a$ and $b$ in $I^*$ such that $s < a < st(a) = st(b) < b < t$ and $f^*(a) = f^*(b)$. Then, there exist infinitesimals $h, k > 0$ such that $a = u - h$ and $b = u+k$, where $u = st(a) = st(b)$. Since $s < u < t$, $f$ is differentiable at $u$ and hence, $0 = f^*(b) - f^*(a) = f^*(u+k) - f^*(u-h) = f^*(u+k) - f(u) + f(u) - f^*(u-h) = kf'(u) + k\,\varepsilon(u,k) + hf'(u) + h\,\varepsilon(u,-h)$. Hence, $f'(u) = -(k\,\varepsilon(u,k) + h\,\varepsilon(u,h))/(h+k)$. Since $h > 0$ and $k > 0$ we have $0 < h/(h+k) < 1$ and $0 < k/(h+k) < 1$, i.e., $f'(u)$ is equal to an infinitesimal. Thus $f'(u) = 0$ since $f'(u)$ is a standard number. This completes the proof of Rolle's theorem.

It is well-known that the mean-value theorem does not generalize to complex-valued or vector-valued functions. The main reason for this is that in those cases there is no direct substitute for the intermediate value theorem. It seems therefore to be important to replace the mean-value theorem by another theorem, preferably of equal power, which will generalize to functions which are not necessarily real-valued. The following theorem is such a theorem. We shall give it for real functions however, but we shall arrange the proof in such a way that it will apply immediately to the more general cases discussed above. Furthermore, the reader can check immediately that the following theorem is at least as important as the mean-value theorem.

THEOREM 10.3. Let $f$ be a real continuous function defined on a bounded and closed interval $I = I(s,t)$, $s < t$. If $f$ is differentiable for all $x$ such that $s < x < t$, then there exists a number $u \in R$ with the following properties: $s < u < t$ and $\left|\frac{f(t)-f(s)}{t-s}\right| \leq |f'(u)|$.

PROOF. We shall first prove the following statement which is analogous to Levy's theorem.

(*) For every $n \geq 3$, there exists an interval $I_n = I(s_n, t_n)$ such that $s < s_n < t_n < t$, $t_n - s_n = \frac{t-s}{n}$ and $\left|\frac{f(t)-f(s)}{t-s}\right| \leq \left|\frac{f(t_n)-f(s_n)}{t_n-s_n}\right|$.

Consider the function $g(x) = f(x+\frac{t-s}{n}) - f(x)$, $s \leq x \leq s + \frac{n-1}{n}(t-s)$, where $n \geq 3$. Then $\frac{f(t) - f(s)}{t-s} = \frac{1}{n}\sum_{k=o}^{n-1} n\left(\frac{g(s+\frac{k}{n}(t-s))}{t-s}\right)$. Hence, we have

$$\left|\frac{f(t)-f(s)}{t-s}\right| \leq \frac{1}{n}\sum_{k=o}^{n-1} n\left|\frac{g(s+\frac{k}{n}(t-s))}{t-s}\right|.$$

From this inequality we may conclude that either we have (i) $\left|\frac{f(t)-f(s)}{t-s}\right| = n\left|\frac{g(s+\frac{k}{n}(t-s))}{t-s}\right|$ for all

$k = 0, 1, 2, \ldots, n-1$ or we have (ii) for some $k$, $0 \leq k \leq n-1$,

$$\left|\frac{f(t)-f(s)}{t-s}\right| < n \left|\frac{g(s+\frac{k}{n}(t-s))}{t-s}\right| .$$ Since $n \geq 3$, there is nothing to prove in case (i). In case (ii), observe that the continuity of $g$ implies that there exists a number $u \in R$ such that $u > s + \frac{k-1}{n}(t-s)$ and

$$n \left|\frac{f(u+\frac{t-s}{n}) - f(u)}{t - s}\right| > \left|\frac{f(t)-f(s)}{t-s}\right| .$$ This proves (*). To complete the proof, observe that, as in the case of the proof of Rolle's theorem, there exist a number $u \in R$ and infinitesimals $h, k > 0$ such that $s < u < t$, and

$$\left|\frac{f(t)-f(s)}{t-s}\right| < \left|\frac{f^*(u+k)-f^*(u-h)}{h-k}\right| = \left|\frac{f^*(u+k)-f(u)+f(u)-f^*(u-h)}{h+k}\right| < |f'(u)|$$
$$+ \frac{k\,|\,\varepsilon(u,k)|+h|\,\varepsilon(u,-h)|}{h+k} = |f'(u)| + \ell .$$ Since $\ell \in M_1$, the required result follows by taking standard parts.

For generalizations of Theorem 10.3 we refer the reader to section 5 of the following text: J. Dieudonné, Foundations of Modern Analysis, New York (1960), Chapter VIII.

## 11. Uniform Convergence; Equicontinuity.

Let $\{f_n : n \in N\}$ be a sequence of real functions defined on a non-empty subset $S$ of $R$. The sequence $\{f_n : n \in N\}$ is said to be convergent at $x \in S$ if $\lim_{n \to \infty} f_n(x)$ exists. If this limit exists for all $x \in S$, then the sequence is said to be <u>convergent</u> on $S$ and the function $f(x)=\lim_{n\to\infty} f_n(x)$ is called the <u>limit</u> of the sequence.

From Theorem 1.1 of Chapter 2 the following theorem follows immediately.

THEOREM 11.1. <u>Let $\{f_n : n\in N\}$ be a sequence of real functions defined on a non-empty point set $S$ of $R$. Then the sequence $\{f_n : n\in N\}$ converges on $S$ if and only if there exists a real function $f(x)$ on $S$ such that $f^*_\omega(x) =_1 f(x)$ for all $x \in S$ and all $\omega \in N^* - N$.</u>

From Cauchy's criterion (Theorem 3.1 of Chapter 2) the following theorem follows immediately.

THEOREM 11.2. <u>Let $\{f_n : n\in N\}$ be a sequence of real functions defined on a non-empty point set $S$ of $R$. Then the sequence $\{f_n : n\in N\}$ converges on $S$ if and only if $f^*_\omega(x) =_1 f^*_{\omega'}(x)$ for all $\omega, \omega' \in N^* - N$ and all $x \in S$. In this case the limit function $f(x)$ of the sequence is given by the formula $f(x) = st(f^*_\omega(x))$, $x \in S$ and $\omega \in N^* - N$.</u>

With these two theorems the elementary rules of the theory of limits of sequences of functions can now easily be obtained. We shall leave this to the reader to verify.

For the discussion of uniform convergence in non-standard analysis, the following definition, which is analogous to Definition 6.1 of Chapter 2, is fundamental.

DEFINITION 11.1 (Non-standard extension of a sequence of real functions) Let $\{f_n : n \in N\}$ be a sequence of real functions defined on a non-empty subset $S$ of $R$. Then the sequence $\{f_n^* : n \in N^*\}$ of functions on $S^*$ defined as follows: $F \in f_n^*(a)$ if and only if $(\exists M)(\exists A)(A \in a,\ M \in n$ and $\{k : F(k) = f_{M(k)}(A(k))\} \in \mathcal{U})$ is called the non-standard extension of the sequence.

It is obvious from the definition that $(\exists M)(\exists A)(A \in a, M \in n$ and $\{h : F(k) = f_{M(k)}(A(k)) \in \mathcal{U}) \Rightarrow (\forall M)(\forall A)(A \in a,\ M \in n \Rightarrow \{h : F(k) = f_{M(k)}(A(k))\} \in \mathcal{U})$. Furthermore the non-standard extension of a sequence of functions defined on a non-empty subset $S$ of $R$ is a non-standard sequence of non-standard functions defined on $S^*$.

Let $\{f_n : n \in N\}$ be a sequence of functions defined on a non-empty point-set $S$ of $R$. Recall that the sequence $\{f_n : n \in N\}$ is said to converge uniformly on $S$ to a function $f(x)$ if for every $\varepsilon \in R_+$ there exists an index $n_o(\varepsilon)$ (which depends on $\varepsilon$, $f$ but not on $x$) such that $|f_n(x) - f(x)| < \varepsilon$ for all $n \geq n_o$ and all $x \in S$. Hence, $f_\omega^*(a) =_1 f^*(a)$ for all $\omega \in N^* - N$ and all $a \in S^*$ is the statement to which the previous statement reduces in non-standard analysis. The converse, however, is also true. Indeed, the following theorem holds.

THEOREM 11.3. Let $\{f_n : n \in N\}$ be a sequence of real functions defined on a non-empty point set $S$ of $R$. Then the sequence $\{f_n : n \in N\}$ converges uniformly to the real function $f(x)$ on $S$ if and only if $f_\omega^*(a) =_1 f^*(a)$ for all $\omega \in N^* - N$ and all $a \in S^*$.

PROOF. We have only to show that the condition of the theorem is sufficient. To this end, we assume that not($\lim f_n(x) = f(x)$ uniformly on S) holds. Then there exist a mapping A of N into S and an injection $\Omega$ of N into N and a positive number $\varepsilon_0 \in R_+$ such that $| f_{\Omega(n)}(A(n)) - f(A(n))| > \varepsilon_0$ for all $n \in N$. Hence, by Definition 11.1, we have $f^*_\omega(a) \neq_1 f^*(a)$, where $\Omega \in \omega$ and $A \in a$. This contradicts the hypothesis and finishes the proof of the theorem.

REMARK. From Theorems 11.1 and 11.3 it follows immediately that if a sequence of functions converges uniformly on some subset of R, then it converges on that subset. Furthermore, by Theorem 7.2 of Chapter 1, on finite sets the two notions of convergence and uniform convergence of sequences of functions coincide.

The well-known theorem that uniform Cauchy sequences of functions converge uniformly takes the following form in non-standard analysis.

THEOREM 11.4. <u>Let</u> $\{f_n : n \in N\}$ <u>be a sequence of real functions defined on a non-empty subset</u> S <u>of</u> R. <u>Then the sequence</u> $\{f_n : n \in N\}$ <u>converges uniformly on</u> S <u>if and only if</u> $f^*_\omega(a) =_1 f^*_{\omega'}(a)$ <u>for all</u> $\omega, \omega' \in N^*-N$ <u>and all</u> $a \in S^*$.

PROOF. That the condition is necessary follows immediately from the preceding theorem. In order to prove that the condition is sufficient, observe that Theorem 11.2 implies that the sequence $\{f_n : n \in N\}$ converges on S to $f(x) = st(f^*_\omega(x))$ where $\omega \in N^*-N$. We shall prove that the convergence is uniform. To this end, observe that $\lim_{n\to\infty} f_n(x) = f(x)$ on $S$ implies that to every $a \in S^*$ there corresponds an index $\omega \in N^*-N$ such that $f^*_\omega(a) =_1 f^*(a)$. Indeed, if $A(n) \in a$ and $\lim_{n\to\infty} \varepsilon_n = 0$, where $\varepsilon_n \in R_+$, then there exists an injection $\Omega$ of $N$ into $N$ such that $|f_{\Omega(n)}(A(n))-f(A(n))| < \varepsilon_n$. Finally, since $f^*_\omega(a) =_1 f^*_{\omega'}(a)$ for all $\omega,\omega' \in N^* - N$ and $a \in S^*$, we obtain that $f^*_\omega(a) =_1 f^*(a)$ for all $\omega \in N^*-N$ and all $a \in S^*$. Then the required result follows from Theorem 11.1.

A concept closely related to that of uniform convergence is that of equicontinuity. Recall that a sequence of continuous real functions $\{f_n : n \in N\}$ defined on a non-empty subset $S$ of $R$ is equicontinuous at $x \in S$ it for every $\varepsilon \in R_+$ there exists a neighborhood $V$ of $x$ such that $|f_n(x)-f_n(y)| < \varepsilon$ for all $n \in N$ and all $y \in V \cap S$. Hence, $f^*_\omega(x) =_1 f^*_\omega(x+h)$ for all $\omega \in N^*-N$ and all $h \in M_1$ such that $x+h \in S^*$, is the statement to which equicontinuity reduces in non-standard analysis. The converse holds also. In fact, we have the following theorem.

THEOREM 11.5. *Let $\{f_n : n \in N\}$ be a sequence of continuous real functions defined on a non-empty point set $S$ of $R$. Then the sequence $\{f_n : n \in N\}$ is equicontinuous at $x \in S$ if and only if $f^*_\omega(x+h) =_1 f^*_\omega(x)$ for all $\omega \in N^*-N$ and all $h \in M_1$ such that $x + h \in S^*$.*

PROOF. We have only to show that the condition is sufficient. To this end, assume that the sequence $\{f_n : n \in N\}$ of real continuous functions is not equicontinuous at $x \in S$. Then there exist an injection $\Omega$ of $N$ into

$N$ and a mapping $A$ of $N$ into $S$ and a positive number $\varepsilon_o \in R_+$ such that $\lim_{n\to\infty} A(n) = x$ and $|f_{\Omega(n)}(A(n))-f_\Omega(x)| > \varepsilon_o$. Hence, by Definition 11.1 we have $f^*_\omega(a) \neq_1 f^*_\omega(x)$, where $a =_1 x$ and $a \in S^*$ and $\omega \in N^*-N$. This contradicts the assumption and finishes the proof of the theorem.

We have the following sufficient condition for equicontinuity.

THEOREM 11.6. <u>Let $\{f_n:n\in N\}$ be a sequence of continuous real functions defined on a non-empty subset</u> $S$ <u>of</u> $R$. <u>If the sequence converges uniformly on</u> $S$, <u>then it is equicontinuous at every point of</u> $S$ <u>and its limit is continuous on</u> $S$.

PROOF. If $\{f_n:n\in N\}$ converges uniformly on $S$, then given $\varepsilon > 0$ there exists an idex $n_o$ such that $m \geq n_o$ implies $|f_{n_o}(x)-f_m(x)| < \varepsilon$ all $x \in S$. Since $f_{n_o}$ is continuous at $x$ we have that $f^*_{n_o}(x+h) =_1 f_n(x)$ for all $h \in M_1$ such that $x \in h \in S^*$. Hence, $|f^*_{n_o}(x+h)-f_m(x)| < \varepsilon$ for all $m \geq n_o$. Thus $|f^*_{n_o}(x+h) - f^*_\omega(x)| \leq \varepsilon$ for all $h \in M_1$ such that $x+h\in S$ and <u>all</u> $\omega\in N^* - N$. It is evident, however, that $|f^*_{n_o}(x+h)-f^*_\omega(x+h)| \leq \varepsilon$. Hence, $f^*_\omega(x) =_1 f^*_\omega(x+h)$ for all $h \in M_1$ such that $x + h \in S^*$ and all $\omega \in N^* - N$. This proves that the sequence is equicontinuous. In order to prove the second part of the theorem let $f(x) = \lim_{n\to\infty} f_n(x)$. Then $f^*(x+h) =_1 f^*_\omega(x+h) =_1 f^*_\omega(x) =_1 f(x)$ for all $h \in M_1$ such that $x + h \in S^*$. This finishes the proof of the theorem.

There is a partial converse to the preceding theorem.

THEOREM 11.7. Let $\{f_n : n \in N\}$ be a sequence of continuous real functions defined on a non-empty bounded and closed subset $S$ of $R$ which converges on $S$. Then the sequence converges uniformly on $S$ if and only if the sequence $\{f_n : n \in N\}$ is equicontinuous.

PROOF. We have only to show that the condition is sufficient since its necessity follows from the preceding theorem. To this end, let $a \in S^*$. Then, since $S$ is bounded and closed, there exist an infinitesimal $h$ and an element $x \in S$ such that $x + h = a$. If $\omega, \omega' \in N^* - N$, then $f^*_\omega(a) = f^*_\omega(x+h) =_1 f^*_\omega(x) =_1 f^*_{\omega'}(x) = f^*_{\omega'}(a)$. Hence, by Theorem 11.4, the sequence converges uniformly on $S$.

An immediate consequence of the preceding theorem is the following theorem.

THEOREM 11.8. (Bendixson) Let $\{f_n : n \in N\}$ be a sequence of real functions defined on a bounded and closed subset $S$ of $R$ which converges on $S$. If there exists a constant $M > 0$ such that $|f_n(x) - f_n(y)| \leq M|x-y|^\alpha$ for all $x, y \in S$ and all $n \in N$, where $0 < \alpha \leq 1$, then the sequence converges uniformly on $S$.

PROOF. Observe that $|f_n(x) - f_n(y)| \leq M\,|x-y|^\alpha$ for all $n \in N$ and all $x, y \in S^*$ implies that $|f^*_n(a) - f^*_n(b)| \leq M|a-b|^\alpha$ for all $a, b \in S^*$ and all $n \in N^*$. Since $S$ is bounded and closed, for every $a \in S^*$ there exist $x \in S$ and $h \in M_1$ such that $x + h = a$. Hence, if $a \in S^*$ and $\omega \in N^* - N$ we obtain that $|f^*_\omega(x+h) - f^*_\omega(x)| \leq M\,|h|^\alpha$, i.e., $f^*_\omega(x+h) =_1 f^*_\omega(x)$ for all $h \in M_1$ such that $x + h \in S^*$ and all $\omega \in N^* - N$. Thus the sequence is equicontinuous and the preceding theorem applies.

REMARK. Bendixson's theorem applies to the case that $S$ is a bounded and closed interval and $|f_n'(x)| \leq M$ for all $n \in N$ and all $x \in S$.

We shall now prove Dini's famous theorem in non-standard analysis.

THEOREM 11.9 (U. Dini). Let $\{f_n : n \in N\}$ be a sequence of continuous real functions defined on a non-empty bounded and closed set $S$ of $R$. If the sequence is a monotone sequence which converges on $S$ to a continuous function, then the sequence converges uniformly on $S$.

PROOF. There is no loss in generality if we assume that the sequence of continuous real functions $\{f_n : n \in N\}$ decreases to zero on $S$. Then, if $a \in S^*$ and $\omega \in N^* - N$ we have $0 \leq f_\omega^*(a) \leq f_n^*(a)$ for all $n \in N$. Since $S$ is bounded and closed we have $x = st(a) \in S$. Hence, if we now take standard parts and use the fact that the functions $f_n (n \in N)$ are continuous on $S$ we obtain that $0 \leq st(f_\omega^*(a)) \leq st(f_n^*(a)) = f_n(x)$ for all $n \in N$, i.e., $f_\omega^*(a) =_1 0$ for all $\omega \in N^* - N$ and all $a \in S^*$. This proves the theorem.

The following concept was introduced by W. H. Young. A sequence $\{f_n : n \in N\}$ of real functions defined on a non-empty subset $S$ of $R$ is said to be uniformly convergent at $x \in S$ if for every $\varepsilon \in R_+$ there exist a neighborhood $V$ of $x$ and an index $n_o(\varepsilon)$ such that $|f_n(y) - f_m(y)| < \varepsilon$ for all $n, m \geq n_o$ and all $y \in V \cap S$. Hence, $f_\omega^*(x+h) =_1 f_{\omega'}^*(x+h)$ for all $\omega, \omega' \in N^* - N$ and all $h \in M_1$ such that $x + h \in S^*$. This suggests the following theorem.

THEOREM 11.10. *Let* $\{f_n : n \in N\}$ *be a sequence of real functions defined on a non-empty subset* $S$ *of* $R$. *Then the sequence converges uniformly at* $x \in S$ *if and only if* $f^*_\omega(x+h) =_1 f^*_{\omega'}(x+h)$ *for all* $\omega, \omega' \in N^*-N$ *and all* $h \in M_1$ *such that* $x+h \in S^*$.

PROOF. We have only to show that the condition is sufficient. If the sequence does not converge uniformly at $x \in S$. Then there exist a mapping $H$ of $N$ into $R$ and injections $\Omega$, $\Omega'$ of $N$ into $N$ and a number $\varepsilon_0 \in R_+$ such that $x + H(n) \in S$, $\lim_{n\to\infty} (x+H(n)) = 0$ and $|f_{\Omega(n)}(x+H(n)) - f_{\Omega'}(x+H(n))| \geq \varepsilon_0$, i.e., $f^*_\omega(x+h) \neq_1 f^*_\omega(x+h)$, where $H \in h$ and $\Omega \in \omega$, $\Omega' \in \omega'$. This is a contradiction and the theorem is proved.

From this theorem the following well-known thecrem follows immediately.

THEOREM 11.12 (*W. H. Young*). *A sequence of real functions* $\{f_n : n \in N\}$ *defined on a non-empty bounded and closed subset* $S$ *of* $R$ *converges uniformly on* $S$ *if and only if it converges uniformly at every point of* $S$.

PROOF. We have only to show that the condition is sufficient its necessity is obvious. To this end, let $a \in S^*$, $\omega$, $\omega' \in N^* - N$. Observing that $st(a) \in S$ since $S$ is bounded and closed we obtain that $f^*_\omega(a) = f^*_\omega(x+h) =_1 f^*_{\omega'}(x+h) = f^*_{\omega'}(a)$ where $x = st(a)$ and $x + h = a$. This completes the proof of the theorem.

Finally we shall briefly discuss the following situation. Let $f(x,t)$ be a real function defined for all $x \in S \subseteq R$ and all $t \in T$. If $s$ is an adherent point of $S$, then we say that $\lim_{S \ni x \to s} f(x,t) = f(s,t) = f(t)$ exists uniformly in $t \in T$ if for every $\varepsilon \in R_+$ there exists a number

$\delta(\varepsilon) \in R_+$ such that $|f(y,t)-f(t)| < \varepsilon$ for all $y \in S$, $y \neq x$ and $|y-x| < \delta$ and all $t \in T$. Hence, $f^*(s+h,t) =_1 f^*(t)$ for all $t \in T^*$ and all $h \in M_1$ such that $h \neq 0$ and $s + h \in S^*$. The converse holds also. The following result is also easy to verify.

THEOREM 11.13. *Let* $f(x,t)$ *be a real function defined for all* $x \in S \subseteq R$ *and all* $t \in T \subseteq R$. *Then* $\lim\limits_{S \ni x \to s} f(x,t)$ *exists uniformly in* $t$, *where* $s$ *is an adherent point of* $S$ *if and only if* $f^*(s+h,t) =_1 f(s+k,t)$ *for all* $t \in T^*$ *and all* $h, k \in M_1$ *such that* $h \neq 0$, $k \neq 0$ *and* $s + h \in S^*$ *and* $s + k \in S^*$.

We conclude this section with the following application.

THEOREM 11.14. *Let* $f(x)$ *be a differentiable real function defined on a bounded and closed interval* $I$ *of* $R$. *Then its derivative* $f'(x)$ *is continuous on* $I$ *if and only if* $\lim\limits_{I \ni y \to x} \dfrac{f(x)-f(y)}{x-y} = f'(x)$ *uniformly in* $x$.

PROOF. If $\lim\limits_{I \ni y \to x} \dfrac{f(x)-f(y)}{x-y} = f'(x)$ is uniform in $x$, then for all $h,k \in M_1$ such that $x + h + k \in I^*$ and $x+k \in I^*$ we have $f^*(x+h+k)-f^*(x+k)$ $=_1 h(f')^*(x+k)$. Since $f^*(x+h+k)-f^*(x+h)-f(x)$ we obtain that $(f')^*(x+k)$ $=_1 f'(x)$. This shows that $f'$ is continuous at $x \in I$. Conversely, assume that $f'$ is continuous on $I$. From the mean-value theorem it follows that if $x \in I$, $h,k \in M_1$ such that $x + h + k \in I^*$ and $x + k \in I^*$, then $f^*(x+h+k) - f^*(x+k) = h(f')^*(x+k+\theta h)$ for some $0 < \theta < 1$. Hence, the continuity of $f'$ implies that $f^*(x+h+k)-f^*(x+k) =_1 h(f')^*(x+k)$. This completes the proof of the theorem.

## 12. The Riemann Integral.

We conclude this chapter with a brief discussion of the notion of Riemann-integrability in non-standard analysis.

There are known several but equivalent forms for the notion of Riemann integrability. For instance, if $f$ is a bounded and real function defined on a bounded and closed interval $I = \{x : s \leq x \leq t\}$ in $R$, then $f$ is Riemann integral over $I$ if and only if there exists a constant $A$ which has the following property: For every $\varepsilon \in R_+$ there exists an index $n_o(\varepsilon) \in N$ such that $\left|\frac{t-s}{n} \sum_{k=1}^{n} f(t_k) - A\right| < \varepsilon$, for all $n \geq N$ and all $t_k$ such that $s + \frac{k-1}{n}(t-s) < t_k \leq s + \frac{k}{n}(t-s)$, $k = 1, 2, \ldots, n$. Hence, $\frac{t-s}{\omega} \sum_{k=1}^{\omega} f^*(t_k^*) =_1 A$ for all $\omega \in N^* - N$ and all $t_k^*$ such that $s + \frac{k-1}{\omega}(t-s) < t_k^* \leq s + \frac{k}{\omega}(t-s)$, $k = 1, 2, \ldots, \omega$. The converse is also true. We have the following theorem.

THEOREM 12.1 <u>A bounded real function</u> $f$ <u>defined on a closed interval</u> $I = \{x : s \leq x \leq t\}$ <u>is Riemann integrable over that interval if and only if there exists a constant</u> $A$ <u>having the following property</u>

$\frac{t-s}{\omega} \sum_{k=o}^{\omega-1} f^*(t_k) =_1 A$ <u>for all</u> $\omega \in N^*-N$ <u>and all</u> $s + \frac{k-1}{\omega}(t-s) \leq t_k \leq s + \frac{k}{\omega}(t-s)$, $k = 1, 2, \ldots, \omega$.

PROOF. The proof follows readily from the preceding remark.

EXAMPLE. If $f$ is Riemann integrable over the interval $s \leq x \leq t$,

$$F(x) = \int_s^x f(u)du = st\left(\frac{1}{\omega} \sum_{k=o}^{\omega-1} f^*\left(s+\frac{k(x-s)}{\omega}\right)\right)(x-s).$$

# CHAPTER 4

# FUNCTIONS OF SEVERAL VARIABLES

## 1. A Non-Standard Model for $R_p$.

As is customary in mathematics we shall denote the p-dimensional ($p \geq 1$) real Euclidean space by $R_p$. We shall denote the elements of $R_p$ by $x, y, \ldots$; and the k-th coordinate of an element $x = (x_1, \ldots, x_p) \in R_p$ shall be denoted by $x_k$. We shall denote the inner product of two elements $x, y \in R_p$ by $(x,y) = \sum_{k=1}^{p} x_k y_k$. The norm of an element $\sqrt{(x,x)} = \sqrt{\sum_{k=1}^{p} x_k^2}$ shall be denoted as usual by $||x||$.

The construction of a non-standard model for $R_p$ ($p \geq 1$) in the form of an ultrapower is identical to the construction in the case $p = 1$. We consider the set $R_p^N$ of <u>all</u> mappings of $N$ into $R_p$. Then we introduce an equivalence relation with respect to the ultrafilter $\mathcal{U}$ in the same way as in the case $p = 1$. The set of all equivalence classes will be denoted by $R_p^*$. The linear space structure of $R_p$ is then carried over to $R_p^*$ with the same procedure as in the case $p = 1$. From this construction we see immediately that to every element $a \in R_p^*$ there corresponds uniquely a p-tuple $(a_1, \ldots, a_p)$ of elements of $R^*$. Hence, the linear space $R_p^*$ is isomorphic to the linear space $R^* \times \cdots \times R^*$ (p-times). We formulate this in the following theorem.

THEOREM 1.1. <u>The linear space</u> $R_p^*$ <u>is isomorphic to the linear space</u> $R^* \times \cdots \times R^*$ (p-<u>times</u>), <u>and hence, is a p-dimensional linear space over</u> $R^*$.

An element $a \in R^*_p$ will be called <u>standard</u> if $x_k \in R$ for <u>all</u> $k=1, 2, \ldots, p$. In the other case, it will be called <u>non-standard</u>.

The following definition is similar to Definition 5.1 of Chapter 1.

DEFINITION 1.1. (<u>Infinitesimals, finite and infinite elements</u>) <u>An element</u> $a \in R^*_p$ <u>is called an infinitesimal if</u> $a_k$ <u>is infinitesimal for all</u> $k=1, 2, \ldots, p$. <u>A number</u> $a \in R^*_p$ <u>is called infinite if for at least one index</u> $k$ $(1 \leq k \leq p)$ <u>the</u> k-th <u>coordinate is infinitely large. In the other case,</u> $a$ <u>is called finite</u>.

The set of all finite elements of $R^*_p$ will be denoted by $M^p_o$ and the set of all infinitesimal will be denoted by $M^p_1$. It is easy to see that $M^p_o$ and $M^p_1$ are linear spaces over $M_o$. Furthermore, $M^p_1$ is a linear subspace of $M^p_o$ and Theorem 5.2 of Chap. 1 implies that $M^p_o/M^p_1$ is isomorphic with $R_p$. We have therefore the following theorem.

THEOREM 1.2. <u>If</u> $M^p_o$ <u>and</u> $M^p_1$ <u>are considered as linear spaces over</u> $R$, <u>then</u> $M^p_1$ <u>is a linear subspace of</u> $M^p_o$ <u>and</u> $M^p_o/M^p_1$ <u>is isomorphic to</u> $R_p$.

The linear space homomorphism of $M^p_o$ onto $R_p$ with kernel $R_p$ will be denoted by "st" and called the <u>standard part</u> homomorphism since it is the natural extension of the homomorphism "st" of $M_o$ onto $R$ to $M^p_o$. In this case we shall also write $a =_1 b$ if $a - b \in M^p_1$ and we shall say that $a$ and $b$ are <u>infinitely close</u>. Thus, <u>every finite element of</u> $R^*_p$ <u>is infinitely close to its standard part</u>.

Since st denotes the linear homomorphism of $M^p_o$ onto $R_p$ with kernel $M^p_1$ the following properties of this homomorphism are evident.

(i) *If* $a \in M_0^p$, *then* $st(a) = (st(a_1,\ldots,st(a_p))$.

(ii) *If* $a,b \in M_0^p$, *then* $st(a+b) = st(a) + st(b)$.

(iii) *If* $a \in M_0^p$ *and* $\lambda \in M_0$, then $st(\lambda a) = st(\lambda)\, st(a)$.

(iv) *If* $a \in M_0^p$, *then* $st(a) = 0$ if and only if $a \in M_1^p$.

(v) *If* $x \in R_p$, *then* $st(x) = x$.

It is evident from the definition of the inner product of two elements of $R_p$ that this bilinear functional extends to $R_p^*$ in a natural way as follows: If $a,b \in R_p^*$, then $\sum_{k=1}^{p} a_k b_k$ *is called the (non-standard) inner product of the elements* $a$ *and* $b$ *and will be denoted by* $(a,b)$. *The norm* $\sqrt{(a,a)}$ *of an element* $a \in R_p^*$ *shall again be denoted by* $||a||$.

We may add the following property to the list of properties of st.

(vi) *If* $a,b \in M_0^p$, *then* $st(a,b) = (st(a), st(b))$. *In particular if* $a \in M_0^p$, *then* $st(||a||) = ||st(a)||$.

The statement $a =_1 b$ *is equivalent to* $||a-b||$ *is infinitesimal*.

Finally, we would like to remark that $R_p^*$ carries a *natural topology* defined as the product topology of the interval topology of $R^*$ (Section 6 of Chap. 1). In this topology $R_p^*$ is *disconnected*. Indeed, $M_1^p$ is both open and closed. Further properties of this topology can be easily deduced from the properties of the interval topology of $R^*$.

## 2. Non-Standard Extensions.

Definition 7.1 of Chap. 1 extends as follows.

DEFINITION 2.1 (*Non-standard extension of a subset of* $R_p$) *Let* $S$ *be a subset of* $R_p$. *The set of all* $a \in R_p^*$ *for which there exists an element* $A \in a$ *such that* $\{n : A(n) \in S\} \in \mathcal{U}$ *is called the non-standard extension of* $S$ *and is denoted by* $S^*$.

As in the case of Definition 7.1 of Chapter 1 we have the following equivalent form of the preceding definition: $a \in S^* \Leftrightarrow (\forall A)(A \in a \Rightarrow \{n : A(n) \in S\} \in \mathcal{U})$. In coordinates this definition reads: $a \in S^* \Leftrightarrow (\forall A)(A \in a \Rightarrow \{n : (A_1(n), \ldots, A_p(n)) \in S\} \in \mathcal{U})$.

With this definition it is easy to see that Theorems 7.1 - 7.4 carry over. One may add, however, the following property:

If $S = S_1 \times \ldots \times S_p$, where $S_1, \ldots, S_p$ are subsets of $R$, then $S^* = S_1^* \times \ldots \times S_p^*$.

Theorem 7.5 of Chapter 1 takes on the following form.

THEOREM 2.1. A subset $S$ of $R_p$ is bounded if and only if $S^* \subseteq M_o^p$.

PROOF. If $S$ is bounded , then there is nothing to prove. Conversely, if $S^* \subseteq M_o^p$, then the set of all norm values of the elements is a subset of $M_o$. Hence, by Theorem 7.5 of Chapter 1 we obtain that this set is bounded. This completes the proof of the theorem.

We shall now discuss the non-standard extension of a binary relation with domain in $R_p$ and range in $R_q$, or equivalently, the non-standard extension of a subset of $R_p \times R_q$.

DEFINITION 2.2 (Non-standard extension of binary relations) Let $\Phi$ be a binary relation with domain in $R_p$ and range in $R_q$. Then the relation $(\exists A)(\exists B)$ ($A \in a$ and $B \in b$ and $\{n : (A(n), B(n)) \in \Phi\} \in \mathcal{U}$) between the elements $a, b$ of $R_p^*$ and $R_q^*$ respectively is called the non-standard extension of $\Phi$ and will be denoted by $\Phi^*$.

In this case we have also that $(\Delta_1\Phi)^* = \Delta_1\Phi^*$ and $(\Delta_2\Phi)^* = \Delta_2\Phi^*$. Hence, *if* $f$ *is a mapping of* $R_p$ *into* $R_q$ *defined on some subset* $S$ *of* $R_p$, *then its non-standard extension* $f^*$ *is a mapping of* $R_p^*$ *into* $R_q^*$ *defined on* $S^*$ *and* $(f(S))^* = f^*(S^*)$.

The definition of the non-standard extension of a double sequence (Definition 6.1 of Chapter 2) is now easily seen to be a particular case of Definition 2.2.

## 3. Convergence of Sequences in $R_p$.

Let $\{x_n : n \in N\}$ be a sequence of elements of $R_p$, i.e., a mapping of $R$ into $R_p$ with domain $N \subseteq R$. Recall that $\{x_n : n \in N\}$ is said to be convergent to an element $x \in R_p$ if and only if for every $\varepsilon \in R_+$ there exists an index $n_o(\varepsilon) \in N$ such that $n \geq n_o(\varepsilon)$ implies $||x_n - x|| < \varepsilon$. Hence, $||x^*_\omega - x|| \in M_1$ for all $\omega \in N^* - N$. Thus we have the following theorem analogous to Theorem 1.1 of Chapter 2.

THEOREM 3.1. *A sequence* $\{x_n : n \in N\}$ *of elements of* $R_p$ *is convergent to an element* $x \in R_p$ *if and only if* $x^*_\omega =_1 x$ *for all* $\omega \in N^* - N$.

Cauchy's criterion takes on the following form.

THEOREM 3.2. *A sequence* $\{x_n : n \in N$ *of elements of* $R_p$ *is convergent if and only if* $x^*_\omega =_1 x^*_{\omega'}$ *for all* $\omega, \omega' \in N^* - N$.

These two theorems can be used to prove the basic properties of convergent sequences in $R_p$.

## 4. The Topology of $R_p$.

In section 1 of Chapter 3 we have given the non-standard forms of the basic topological notions and properties of $R$. This can be done in exactly the same way for $R_p$. Therefore we shall merely collect a list of those statements in the following theorem. The proof is left to the reader.

THEOREM 4.1 *Let $S$ be a subset of $R_p$.*

(i) *If $s \in R_p$, then $s \in \bar{S}$, the closure of $S$ in $R_p$, if and only if there exists an element $a \in S^*$ such that $s = st(a)$*

(ii) *An element $s \in R_p$ is a non-trivial adherent point of $S$ if and only if there exists an element $h \in M_1^p$ such that $h \neq o$ and $s + h \in S^*$*

(iii) *An element $s \in R_p$ is a trivial adherent point of $S$, i.e., an isolated point of $S$, if and only if $s + h \notin S^*$ for all $h \in M_1^p$ such that $h \neq o$.*

(iv) *An element $s \in R_p$ is an interior point of $S$ if and only if $s + h \in S^*$ for all $h \in M_1^p$.*

(v) *An element $s \in R_p$ is an exterior point of $S$ if and only if $s + h \notin S^*$ for all $h \in M_1^p$.*

The following proof of the Bolzano-Weierstrass theorem is based on Theorem 2.1 and (ii) of Theorem 4.1

THEOREM 4.1 (*Bolzano-Weierstrass*). *If $S$ is a bounded infinite subset of $R_p$, then $S$ has at least one non-trivial adherent point.*

PROOF. Since $S$ is infinite we have $S^* \neq S$ (Theorem 7.4 of Chapter 1). By Theorem 2.1, $S$ is bounded implies $S^* \subseteq M_o^p$. Let $a \in S^* - S$. Then $a \in M_o^p$ and hence, by (ii) of Theorem 4.1, $st(a)$ is a non-trivial adherent point of $S$.

5. Limits of Functions.

Let $f$ be a mapping of a non-empty subset $S$ of $R_p$ into $R_q$. If $s \in R_p$ is an adherent point of $S$, then we say that $f(x)$ converges to $\ell \in R_q$ if $x$ tends to $s$ through $S$ if and only if $(\forall \varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists \delta)(\delta \in R_+$ and $(\forall x)(x \in S$ and $0 < \|x-s\| < \delta \Rightarrow \|f(x)-\ell\| < \varepsilon)))$. In that case we write $\lim\limits_{S \ni x \to s} f(x) = \ell$. Hence, we have the following therem analogous to Theorem 2.1 of Chapter 3.

THEOREM 5.1. *Let $S$ be a non-empty subset of $R_p$ and let $f$ be a mapping of $S$ into $R_q$. If $s \in R_p$ is an adherent point of $S$, then $\lim\limits_{S \ni x \to s} f(x) = \ell$ if and only if $f^*(s+h) =_1 \ell$ for all $h \in M_1^p$ such that $h \neq o$ and $s+h \in S^*$.*

Cauchy's criterion takes on the following form.

THEOREM 5.2 (*Cauchy's criterion*) *Let $S$ be a non-empty subset of $R_p$ and let $f$ be a mapping of $S$ into $R_q$. If $s \in R_p$ is an adherent point of $S$, then $\lim\limits_{S \ni x \to s} f(x)$ exists if and only if $f^*(s+h) =_1 f^*(s+k)$ for all $h, k \in M_1^p$ such that $h \neq o$, $k \neq o$, $s + h \in S^*$ and $s + k \in S^*$.*

6. Continuity

Analogous to Theorem 3.1 of Chapter 3 we have

THEOREM 6.1. *Let $S$ be a non-empty subset of $R_p$ and let $f$ be a mapping of $S$ into $R_q$. Then $f$ is continuous at $s \in S$ if and only if $f^*(s+h) =_1 f^*(s)$ for all $h \in M_1^p$ such that $s+h \in S^*$, or equivalently, $f^*(a) =_1 f^*(b)$ for all $a, b \in S^*$ such that $st(a) = st(b) = s$.*

It is now easy to see that all the results obtained in section 3 of Chapter 3 generalize to functions of several variables.

We conclude this section with the following generalization of the theorem of Weierstrass (Theorem 4.2 of Chapter 3).

THEOREM 6.2 (Weierstrass). *Let* $S$ *be a non-empty subset of* $R_p$ *and let* $f$ *be a mapping of* $S$ *into* $R_q$. *If* $f$ *is continuous on* $S$ *and* $S$ *is bounded and closed, then* $f(S)$ *is bounded and closed.*

PROOF. Let $a \in (f(S))^*$. Since $(f(S))^* = f^*(S^*)$, there exists an element $b \in S^*$ such that $a = f^*(b)$. Hence, $S$ being bounded and closed, there exists an element $s \in S$ and an element $h \in M_1^p$ such that $b = s+h$. Thus, $f$ being continuous, $f^*(b) =_1 f(s)$, i.e., $a$ is finite, or equivalently, $f(S)$ is bounded. In order to prove that $f(S)$ is closed, let $s \in \overline{f(S)}$. Then there exists an element $h \in M_1^p$ such that $s+h \in (f(S))^* = f^*(S^*)$. Hence, there exists an element $a \in S^*$ such that $f^*(a) = s+h$. We conclude that $f(st(a)) = s$, i.e., $s \in f(S)$, or equivalently, $f(S)$ is closed.

## 7. Uniform Continuity.

Let $S$ be a non-empty subset of $R_p$ and let $f$ be a mapping of $S$ into $R_q$. Recall that $f$ is said to be uniformly continuous on $S$ if the following statement holds: $(\forall\varepsilon)(\varepsilon \in R_+ \Rightarrow (\exists\delta)(\delta \in R_+$ and $(\forall x)(\forall y)(x,y \in S$ and $\|x-y\| < \delta \Rightarrow \|f(x) - f(y)\| < \varepsilon)))$. Thus the following theorem holds.

THEOREM 7.1. *Let* $S$ *be a non-empty subset of* $R_p$ *and let* $f$ *be a mapping of* $S$ *into* $R_q$. *Then* $f$ *is uniformly continuous on* $S$ *if and only if* $f^*(a) =_1 f^*(b)$ *for all* $a, b \in S^*$ *such that* $a =_1 b$.

The proof of this theorem is left to the reader since it is similar to the proof of Theorem 6.1 of Chapter 3.

A mapping $f$ of $S \subseteq R_p$ into $R_q$ is continuous on $S$ according to Theorem 6.1 if and only if $f^*(a) =_1 f^*(b)$ for all $a, b \in S^*$ such that $st(a) = st(b) \in S$. If we compare this condition with the condition of Theorem 7.1, then the following theorem is completely evident

THEOREM 7.2 (<u>Heine</u>) <u>Let</u> $S$ <u>be a non-empty subset of</u> $R_p$ <u>and let</u> $f$ <u>be a mapping of</u> $S$ <u>into</u> $R_q$. <u>If</u> $f$ <u>is continuous on</u> $S$ <u>and if</u> $S$ <u>is bounded and closed, then</u> $f$ <u>is uniformly continuous on</u> $S$.

PROOF. Indeed, in this case the condition of being continuous is the same as being uniformly continuous.

## 8. Differentiation of Functions of Several Variables.

Let $S$ be a non-empty subset of $R_p$ and let $f$ be a mapping of $S$ into $R_q$. Recall that, if $x \in S$ is a non-trivial adherent point of $S$, then $f$ is called differentiable at $x \in S$ if there exists a linear transformation $A(x)$ of $R_p$ into $R_q$ such that $f(x+h)-f(x) = A(x)h + ||h|| \, \varepsilon(x,h)$ for all $h \in R_p$ such that $x+h \in S$ and $\lim \varepsilon(x,h) = o$ as $||h|| \to 0, x+h \in S$. In this case, the linear transformation $A(x)$ is called the derivative of $f$ at $x$. As is customary in mathematics we shall use the notation $f'(x)$ for $A(x)$.

Since $x \in S$ is a non-trivial adherant point of $S$ it is easy to see that the linear transformation $f'(x)$ is uniquely determined whenever it exists.

In non-standard analysis this definition becomes.

THEOREM 8.1. <u>Let</u> $S$ <u>be a non-empty subset of</u> $R_p$ <u>and let</u> $f$ <u>be a mapping of</u> $S$ <u>into</u> $R_q$. <u>If</u> $x \in S$ <u>is a non-trivial adherent point of</u> $S$, <u>then</u> $f$ <u>is differentiable at</u> $x \in S$ <u>if and only if there exists a linear transformation of</u> $R_p$ <u>into</u> $R_q$, <u>which we denote by</u> $f'(x)$, <u>such that for all</u> $h \in M_1^p$, $x+h \in S^*$ <u>we have</u>

$$(*) \qquad f^*(x+h) - f(x) = f'(x)h + ||h|| \; \varepsilon(x,h),$$

<u>where</u> $\varepsilon(x,h) =_1 0$.

The proof of this theorem is left to the reader.

Observe that the right-hand side of $(*)$ is infinitesimal. Hence, the following theorem holds.

THEOREM 8.2 <u>At those points where a function is differentiable it is also continuous.</u>

If we denote by $e_1,\ldots,e_p$ the unit vectors $e_k = (0,\ldots,1,\ldots 0)$, $k = 1,2,\ldots,p$, then we see immediately that if $h \in M_1$ and $(*)$ holds at $x$ we have $f^*(x+h\, e_k) - f(x) = h\, f'(x)\; e_k + |h| \;\varepsilon(x,he_k)$. Hence, if $f$ is differentiable at $x$, then $f$ is partially differentiable at $x$ with derivatives $f'(x)e_k$. If we denote the function $f$ in coordinates, i.e., $f = (f_1,\ldots,f_p)$, then $f'(x)e_k$ is the matrix $(\frac{\partial f_i}{\partial x_k}(x))$, $i = 1,2,\ldots,q$. The linear transformation $f'(x)$ is given in coordinates by the matrix

$(\frac{\partial f_i}{\partial x_k})$ $i = 1,2,\dots,q$; $k = 1,2,\dots,p$. This matrix is often referred to as the functional matrix. If $p = q$, the determinant of the functional matrix of $f$ is called the Jacobian of $f$.

We shall now prove the chain rule.

THEOREM 8.3. Let $f$ be a mapping of $S \subseteq R_p$ into $R_q$ and let $g$ be a mapping of $T \subseteq R_q$ into $R_\ell$ such that $f(S) \subseteq T$. If $x \in S$ and $x$ is a non-trivial adherent point of $S$ and if $f(x)$ is a non-trivial adherent point of $T$, then $f$ differentiable at $x$ and $g$ differentiable at $f(x)$ implies that the composite function $g \circ f = g(f)$ is differentiable at $x$ and $g(f)'(x) = g'(f(x))f'(x)$.

PROOF. Let $0 \neq h \in M_1^p$ such that $x+h \in S^*$. Set $f^*(x+h) - f(x) = k$. Then $k \in M_1^q$. Since $g^*(f^*(x+h)) = g(f)^*(x+h)$, we have $(g(f))^*(x+h) - g(f(x))$ $= g^*(f^*(x+h)) - g(f(x)) = g^*(f(x)+k) - g(f(x)) = g'(f(x))k + ||k||\ \varepsilon_g\ (f(x),k)$. Observe that $k = f'(x)h + ||h||\ \varepsilon_f\ (x,h)$. Hence, $g(f)^*(x+h) - g(f(x))$ $= g'(f(x))\ f'(x)h + ||h||\ g'(f(x))\ \varepsilon_f + \ ||k||\ \varepsilon_g\ (f(x),k) = g'(f(x))f'(x)h$ $+ ||h||\ (g'(f(x))\ \varepsilon_f\ (x,h) + f'(x)\ (\frac{h}{||h||}\ \ \varepsilon_g\ (f(x),k) + \varepsilon_f\ (x,h)\ \varepsilon_g(f(x),k))$. It is easy to see that $f'(x)\ (\frac{h}{||h||})$ is finite; hence the expression between the brackets is infinitesimal. This completes the proof of the theorem.

## 9. The Mean-Value Theorem

We shall now give a formulation of the mean-value theorem for functions of several variables. Observe the similarity between the following theorem and Theorem 10.3 of Chapter 3.

THEOREM 9.1. (*Mean-value theorem*) *Let* $f$ *be a mapping of a non-empty open subset* $S$ *of* $R_p$ *into* $R_q$. *If* $x$ *and* $y$ *are two elements of* $S$ *such that the line segment* $x + t(y-x)$, $0 \leq t \leq 1$, *is in* $S$, *then the following conditions:*

(i) $f(x+t(y-x))$ *is continuous for all* $0 \leq t \leq 1$,

(ii) $f$ *is differentiable at every point* $x + t(y-x)$ *whenever* $0 < t < 1$, *imply that there exists a number* $\theta$ *such that* $0 < \theta < 1$ *and* $||f(y) - f(x)|| \leq ||y-x|| \cdot ||f'(x+\theta(y-x))||$.

PROOF. Set $g(t) = f(x+t(y-x))$, $0 \leq t \leq 1$. Then $g$ is a continuous mapping of $0 \leq t \leq 1$ into $R_q$. Furthermore, $g$ is differentiable for all $0 < t < 1$ and $g'(t) = f'(x+t(y-x))(y-x)$. We shall prove first the following statement .

(L) *For every* $3 \leq n \in N$ *there exist numbers* $s_n$, $t_n$ *such that* $0 < s_n < t_n < 1$, $t_n - s_n = \frac{1}{n}$ *and* $||g(1) - g(0)|| \leq n \, ||g(t_n) - g(s_n)||$.

In order to prove (L), set $h(t) = g(t+\frac{1}{n}) - g(t)$, $0 \leq t \leq 1 - \frac{1}{n}$. Then

$||g(1) - g(0)|| = \left|\left|\sum_{k=0}^{n-1} h(\frac{k}{n})\right|\right| \leq \frac{1}{n} \sum_{k=0}^{n-1} n||h(\frac{k}{n})||$. Hence, we have either

(i) $||g(1)-g(0)|| = n \, ||h(\frac{k}{n})||$ for all $0 \leq k \leq n-1$ or (ii) for some $0 \leq k \leq n-1$ we have $||g(1) - g(0)|| < n \, ||h(\frac{k}{n})||$. Since $n \geq 3$, there is nothing to prove if (i) holds. In case (ii), observe that the continuity of $g$ implies that $||h(\frac{k}{n})|| = ||g(\frac{k+1}{n}) - g(\frac{k}{n})|| = \lim ||g(u+\frac{1}{n}) - g(u)||$ as $u$ tends to $\frac{k}{n}$ and $0 < u < 1$. Hence, for some $u$, $0 < u < 1$ we have $n||g(u+\frac{1}{n})-g(u)|| > ||g(1) - g(0)||$. This completes the proof of (L).

To complete the proof, we apply (L) successively. Then we obtain that there exist a number $\theta$, $0 < \theta < 1$, and infinitesimals $h$, $k$ such that $h, k > 0$ and

$$||g(1) - g(0)|| < \frac{||g^*(\theta+k)-g^*(\theta-h)||}{h+k} = \frac{||g^*(\theta+k)-g(\theta)+g(\theta)-g^*(\theta-h)||}{h+k} =$$

$$= \frac{||g'(\theta)(h+k) + k\varepsilon_g(\theta,k) + h\varepsilon_g(\theta,h)||}{h+k} \leq ||g'(\theta)|| + \ell, \text{ where } \ell \in M_1.$$

If we then take standard parts we obtain $||f(y) - f(x)|| \leq ||g'(\theta)|| = ||f'(x+\theta(y-x))(y-x)|| \leq ||f'(x+\theta(y-x))|| \; ||y-x||$. This completes the proof of the theorem.

# CHAPTER 5

# THE THEORY OF DISTRIBUTIONS

## 1. Quasi-Standard Functions.

In the preceding chapters we have only studied the properties of standard functions in non-standard analysis. For a discussion in non-standard analysis of the theory of distributions, where one does not operate with standard functions, it is necessary to employ more general functions. In the following definition we shall single out a class of functions, which will yield a natural realization of generalized functions or distributions.

DEFINITION 1.1 (Quasi-standard function) Let $S$ be a standard set. A mapping $f$ of $S^*$ into $R^*$ is called a quasi-standard function if there exist a family of standard functions $g_t(x) = g(x,t)$ defined on $S$ for all values of the parameter $t \in T \subseteq R$ and an element $a \in T^*$ such that $f(x) = g^*(x,a)$ for all $x \in S^*$.

It is evident that the non-standard extension of a standard function $f$ is a quasi-standard function. Indeed, let $g(x,t)$ be the family of standard functions which consists of the standard function $f$ only. The converse need not be true as will be shown in the examples below.

EXAMPLES 1. If $\omega \in N^*-N$, then $x^\omega$ is a quasi-standard function defined for all $x \in R^*$. Indeed, consider the family of standard functions $g(x,n) = x^n$, $x \in R$ and $n \in N$. Then $x^\omega = g^*(x,\omega)$, $x \in R^*$ and $\omega \in N^*-N$. It is obvious that $x^\omega$ is not the non-standard extension of a standard function. Indeed, $(\frac{1}{2})^\omega \in M_1$.

2. Let $g = g(x,y)$ be a standard function of two real variables defined in the Cartesian product $S \times T$ of two standard sets $S$ and $T$. Then the function $f(x) = g^*(x,a)$, where $a \in T^*$ is a quasi-standard function defined on $S^*$.

3. (<u>Quasi-standard polynominal</u>) If $p_n(x) = \sum_{k=0}^{n} a_k x^k$ is a sequence of polynomials, then for every $\omega \in N^*-N$, $p^*_\omega(x) = \sum_{k=0}^{\omega} a^*_k x^k$ is a quasi-standard function defined on $R^*$, which we shall call a <u>quasi-standard polynomial of degree</u> $\omega$.

More generally, if $p_n(x) (n \in N)$ is an arbitrary family of polynomials, then for every $\omega \in N^*-N$, the quasi-standard function $p^*_\omega(x)$ $(x \in R^*)$ shall also be called a quasi-standard polynomial. Observe, that in this case the degree of $p^*_\omega$ may be different from $\omega$.

4. (<u>Quasi-standard rational functions</u>). Let $\{r_n, n \in N\}$ be a sequence of rational functions defined in some standard set $S$ of $R$. Then for all $\omega \in N^*-N$, $r = r^*_\omega$ is called a <u>quasi-standard rational function</u>. For instance, $r(x) = 1/(1 + x^2)^\omega$, where $\omega \in N^*-N$ is a quasi-standard rational function defined for all $x \in R^*$.

5. Consider the family of standard functions $\{e^{-nx} : n \in N\}$ defined for all $x \in R$. Then for every $\omega \in N^*-N$, $e^{-\omega x}$ is a quasi-standard function defined for all $x \in R^*$. It is easy to see that $e^{-\omega x}$ is infinitesimal for all standard $x > 0$, and that $e^{-\omega x}$ is infinitely large for all standard $x < 0$.

In the theory of quasi-standard functions it is important to know the relation which may exist between two families of standard functions which generate the same quasi-standard function. The following theorem deals with this question.

THEOREM 1.1 <u>Let</u> $S$ <u>be a standard set</u>. <u>Then a quasi-standard function,</u> $f$ <u>defined on</u> $S^*$ <u>is obtained from two families</u> $\{g_1(x,t) : x \in S$ <u>and</u> $t \in T_1\}$.

$\{g_2(x,t) : x \in S$ and $t \in T_2\}$ of standard functions, i.e., there exist elements $a_1 \in T_1^*$ and $a_2 \in T_2^*$ such that $f(x) = \overset{*}{g}_1(x,a_1) = \overset{*}{g}_2(x,a_2)$ for all $x \in S^*$, if and only if $(\forall A_1)(\forall A_2)$ $(A_1 \in a_1$ and $A_2 \in a_2 \Longrightarrow \{n: g_1(x,A_1(n)) = g_2(x,A_2(n))$ for all $x \in S\} \in \mathcal{U}$ ).

PROOF. If $x \in S^*$, the condition of the theorem implies that for all $X \in x$, $A_1 \in a_1$ and $A_2 \in a_2$ we have, $\{n: g_1(X(n),A_1(n)) = g_2(X(n),A_2(n))\} \in \mathcal{U}$. Hence, $\overset{*}{g}_1(x,a_1) = \overset{*}{g}(x,a_2)$, which proves that the condition of the theorem is sufficient. In order to show that the condition of the theorem is necessary we shall assume that it does not hold. Then there exist elements $A_1 \in a_1$, $A_2 \in a_2$ such that the set E of all $n \in N$ for which the functions $g_1(x,A_1(n))$ and $g_2(x,A_2(n))$ coincide on S does not belong to $\mathcal{U}$ . We conclude, since $\mathcal{U}$ is an ultrafilter, that its complement $F = N-E$ belongs to $\mathcal{U}$ Let $S_n = \{x: g_1(x,A_1(n)) \neq g_2(x,A_2(n))\}$. Then for all $n \in F$ we have that $S_n \neq \emptyset$. Let $X(n)$ be a mapping of N into S such that $X(n) \in S_n$ for all $n \in F$ (apply the axiom of choice). Then for all $n \in F$ we have $g_1(X(n),A_1(n)) \neq g_2(X(n),A_2(n))$. Since $F \in \mathcal{U}$, we conclude finally that $\overset{*}{g}_1(x,a_1) \neq \overset{*}{g}_2(x,a_2)$, where $X \in x \in S^*$. This contradicts the assumption and finishes the proof of the theorem.

We shall now show that all conditions or operators which apply to standard functions can be extended to quasi-standard functions.

DEFINITION 1.2 (Continuous, differential, ..., quasi-standard function). Let S be a standard set and let f be a quasi-standard function defined on $S^*$. If $\pi$ is a property which applies to standard functions (e.g., continuous, differentiable at a point, integrable, etc.), then f is said to have the property $\pi$ if and only if there exists a family of standard functions $\{g(x,t) : x \in S, t \in T \subset R\}$ and an element $a \in T^*$ such that $(f(x) = \overset{*}{g}(x,a)$

for all $x \in S^*$ and there exists an element $A \in a$ with the property that $\{n: g(\cdot,A(n))$ satisfies $\pi\} \in \mathcal{U}$.

It is evident from the definition that $\{n: g(\cdot,A(n))$ satisfies $\pi\} \in \mathcal{U}$ for all $A \in a$. Furthermore, the preceding theorem implies that the same holds for every family of standard function which defines f.

We shall supplement this definition with the following definition.

DEFINITION 1.3 (Functionals of quasi-standard functions). Let S be a standard set and let f be a quasi-standard function defined on $S^*$. If L is an operator or functional which applies to standard functions (e.g., the derivative, integral, etc. $\cdots$), then L is said to apply to f if and only if there exists a family of standard functions $\{g(x,t) : x \in S$ and $t \in T \subseteq R\}$ and an element $a \in T^*$ such that $f(x) = \overset{*}{g}(x,a)$ for all $x \in S^*$ and $\{n: g(\ ,A(n)) \in$ domain of $L\} \in \mathcal{U}$. In that case, L(f) is defined to be $F^*(a)$, where $F(t) = L(g(\ ,t))$ for all $t \in T$ for which it is defined.

In order to justify the last part of the definition we have to show that L(f) is uniquely determined if L applies. For this purpose assume that f is obtained from two families $\{g_1(x,t) : x \in S,\ t \in T_1 \subseteq R\}$ and $\{g_2(x,t) : x \in S,\ t \in T_2 \subseteq R\}$, so that $f(x) = g_1^*(x,a_1) = g_2^*(x,a_2)$ for all $x \in S^*$. Then, if $F_1(t) = L(g_1(\cdot,t))$ and $F_2(t) = L(g_2(\cdot,t))$, Theorem 1.1 implies that $F_1^*(a_1) = F_2^*(a_2)$, which completes the proof.

EXAMPLES 1. Let $f(x) = g^*(x,a)$, $x \in S^*$ and $a \in T^*$, be a quasi-standard function defined on the non-standard extension $S^*$ of a standard set S. If x is a non-trivial adherent point of S and if f is differentiable at $x \in S$, then $f'(x) = h^*(a)$, where $h(t) = g_x'(x,t)$ for all $t \in T$ for which it exists. It is easy to see that all elementary properties of the derivative extend to differentiable quasi-standard functions. For instance, the mean-value theorem becomes: Let f be a differentiable quasi-standard function defined on the non-standard extension $I^*$ of a standard interval $I(u,v)(u<v)$. Then if $p\in I^*$, $q\in I^*$ and $p < q$, there exists an element $\ell \in I^*$ such that $p < \ell < q$ and $f(q) - f(p) = (q-p)f'(\ell)$.

Another fact worth noticing is that a differentiable quasi-standard function is continuous.

2. Let $f(x) = g^*(x,a)$, $x\in I^*$ and $a\in T^*$, be a quasi-standard function defined on the non-standard extension $I^*$ of a standard interval $I = I(u,v)$ $(u < v)$. If f is integrable over I in some definite sense (e.g., Riemann integrable), then $\int_I f(x)ds = h^*(a)$, where $h(t) = \int_u^v g(x,t)dx$ for all $t\in T$ for which it exists. Again the reader won't have any difficulty to find out what properties of the theory of integration extends to integrable quasi-standard functions. It is worth noticing, however, that if the quasi-standard function f is integrable over I and f as well as $\int_I fdx$ are finite, then, in general, $st(\int_I fdx) \neq \int_I st(f)dx$. Indeed, let $g(x)$ be a continuous function which vanishes outside the bounded and closed interval $1 \leq x \leq 2$ and $\int_{-\infty}^{+\infty} g(x)dx = 1$. Consider then the family of functions $g(x,t) = tg(xt)$, for all $t\in R$. If a is an infinitely large positive number, then $f(x) = ag^*(xa)$ in a quasi-standard function defined for all $x\in R^*$ which has the following properties: $f(x) = 0$ for all $x\in R$ and $\int_{-\infty}^{+\infty} f(x)dx = 1$. Hence, in this case, $st\ (\int_{-\infty}^{+\infty} f(x)dx) = 1$ but $\int_{-\infty}^{+\infty} st(f)dx = 0$. The formula $st(\int_I fdx) = \int_I st(f)dx$ holds, however, whenever there exists a family of standard functions which defines f and which are uniformly integrable, i.e., one can pass to the limit under the integral sign.

If f is a quasi-standard function f defined on the non-standard extension $I^*$ of a standard interval I, then we can define its integral over any non-standard subinterval of $I^*$ in the following way: Let $F(p,q,t) = \int_p^q g(x,t)dx$, where $p,q \in I$. Then, if $c,d \in I^*$ we define $\int_c^d f(x)dx = F^*(c,d,a)$. It is easy to see that this definition is independent from the family of standard functions which defines f.

REMARK. Quasi-standard functions defined on the non-standard extension $S^*$ of S can be <u>added</u>, <u>multiplied</u> and <u>divided</u> by one another provided the divisor does not vanish. Observe, however, that the sum, product or quotient of two quasi-standard functions is in general not a quasi-standard function in the sense of Definition 1.1. If, on the other hand, we allow the parameter set T in Definition 1.1. to be a subset of $R_p$ for some p, where p may be different for different quasi-standard functions, then it is easy to see that, indeed, the set of all quasi-standard functions on $S^*$ is a <u>ring</u>.

As we remarked earlier, the properties of standard functions to the extent that they can be formulated in the lower predicate calculus extend to quasi-standard functions. We shall illustrate this by one more example in proving a generalization of the Weierstrass approximation theorem.

THEOREM 1.2 (<u>Weierstrass</u>) <u>Let</u> f <u>be a quasi-standard function defined on the non-standard extension</u> $I^*$ <u>of a bounded and closed standard interval</u> I. <u>If</u> f <u>is continuous</u>, <u>then for every infinitesimal</u> h <u>there exists a quasi-standard polynomial</u> p <u>on</u> $I^*$ <u>such that</u> $|f(x) - p(x)| < h$ <u>for all</u> $x \in I^*$.

PROOF. Let $f(x) = g^*(x,a)$, where $g(x,t)$, $x \in I$ and $t \in T$ is a family of standard functions which defines f. Since f is continuous, $g(x,A(n))$, where $A \in a$, is a continuous function on $I$ for all $n \in E \in \mathcal{U}$. If $h > 0$ is an infinitesimal and $H \in h$, then it follows from the approximation theorem of Weierstrass that for every $n \in E$ there exists a polynomial $p(x,A(n))$ such that $|g(x,A(n)) - p(x,A(n))| < H(n)$ for all $x \in I$. Hence, $|f(x) - p^*(x,a)| < h$ for all $x \in I^*$. This completes the proof of the theorem.

## 2. Dirac Delta-Functions

We pointed out before that quasi-standard functions yield a natural realization of generalized functions. In this section we shall discuss this by means of the Dirac delta-function. Delta-functions can be introduced at various levels of generality. Intuitively speaking, a delta-function defined on an interval $I$ concentrated at $x_0 \in I$ is a function with the properties that $\delta$ is infinitesimal for all standard $x \in I$ and $x \neq x_0$ and $\int_I \delta(x)dx = 1$. It is easy to give examples to show that there exist quasi-standard functions which have the properties of a delta-function. Indeed, consider, for instance, the family of functions $\{\sqrt{\frac{n}{\pi}}\, e^{-n(x-x_0)^2} : n \in N\}$, then for every $\omega \in N^* - N$, the quasi-standard function $\sqrt{\frac{\omega}{\pi}}\, e^{-\omega(x-x_0)^2}$ is a Dirac delta-function defined on $I = R$ and which is concentrated at $x = x_0$.

It is clear that for a given standard interval $I$ and $x_0 \in I$ there are many delta-functions as opposed in the theory of distributions. We shall see, however, in the next section that we are able to develop the theory of distributions, given recently by Mikusinski and Sikorski, within the theory of quasi-standard functions.

We shall introduce now a more general and more precise definition of a delta-function. This definition is similar to the one given by A. Erdélyi in [6].

DEFINITION 2.1 (*Absolutely integrable delta-function*) *Let* I *be a standard interval and let* $x_o \epsilon I$. *A quasi-standard function* $\delta$ *defined on the non-standard extension* $I^*$ *of* I *is called an absolutely integrable delta-function concentrated at* $x_o$ *if it satisfies the following conditions*:

(i) $\delta$ *is continuous on* $I^*$.

(ii) $\int_{I^*} |\delta(x)|dx$ *is finite and* $\int_{I^*} \delta(x)dx = 1$.

(iii) *There exists an infinitesimal interval* J *such that* $x_o \epsilon J C I^*$, $\delta(x) \geqq 0$ *on* J, *and* $\int_{I^*-J} |\delta(x)|dx$ *is infinitesimal*.

We shall show that there are many delta-functions in the sense of this definition.

THEOREM 2.1. *Let* $\gamma(x)$ *be a continuous standard function defined for all* $x \epsilon R$ *such that* $\gamma(x) \geqq 0$ *and* $\int_R \gamma(x)dx = 1$. *Then for every positive infinitely large number* $a \epsilon R^*$, $a\ \gamma^*(ax)$ *is an absolutely integrable Dirac delta-function concentrated at* $x = 0$.

PROOF. It is only necessary to prove that $\delta$ satisfies condition (iii) of Definition 2.1. To this end, let $b = \sqrt{1/a}$ . Then $b \epsilon M_1$ and $\int_{x \geqq b} \delta(x)dx = \int_{x \geqq ba} \gamma(x)dx =_1 0$. Similarly, we have that $\int_{x \leqq -b} \delta(x)dx =_1 0$. Hence, if we denote by J the infinitesimal interval $\{x\colon x \epsilon R^* \text{ and } |x| \leqq b\}$, then $0 \epsilon J$, $\delta \geqq 0$ on J and $\int_{R^*-J} \delta(x)dx =_1 0$, i.e., $\delta$ satisfies (iii) of Definition 2.1. This completes the proof of the theorem.

Delta functions as defined in Definition 2.1 may vanish for all standard values of its argument. Indeed, let $\gamma(x) = 6(x-1)(x-2)$ whenever $1 \leqq x \leqq 2$ and let $\gamma(x) = 0$ otherwise. Then for every positive infinitely large number a, $a\gamma^*(ax)$ is a Dirac delta function in the sense of Definition 2.1. In this case, however, we have that $st(a\gamma^*(ax)) = 0$ for all $x \epsilon R$. As a complement to Theorem 2.1 we have the following theorem.

THEOREM 2.2. *Let* $\gamma(x)$ *be a continuous standard function defined for all* $x \in R$ *such that* $\gamma(x) > 0$ *for all* $x \in R$, $\gamma$ *is even and* $\gamma(x)$ *decreases to zero as* $x$ *tends to infinity and* $\int_R \gamma(x)dx = 1$. *Then for every positive infinitely large number* $a \in R^*$, $a\gamma^*(ax)$ *is an absolutely integrable Dirac delta function concentrated at* $x = 0$ *with the property that* $0 < a\gamma^*(ax) = {}_1 0$ *for all* $0 \neq x \in R$ *and* $a\gamma^*(ax)$ *in infinitely large at* $x = 0$.

PROOF. Observe that $a\gamma^*(ax) > 0$ for all $x \in R$. Furthermore, $\gamma$ being integrable and decreasing as $x$ tends to infinity we have that $x\gamma(x)$ tends to zero as $x$ tends to infinity. Hence, $a\gamma^*(ax) = {}_1 0$ for all $x \in R$ and $x \neq 0$ and is obviously infinitely large at $x = 0$, since $\gamma(0) \neq 0$. This completes the proof of the theorem, since the other properties follow from the preceding theorem.

The most important property of the delta function is the reproducing property which asserts that under certain conditions on $f$, $\int_{I^*} f(x)\delta(x)dx = {}_1 f(c)$, whenever $\delta$ is a delta function on $I$ concentrated at $c \in I$. For the type of delta function we have introduced we shall prove the following theorem.

THEOREM 2.3. *Let* $\delta$ *be an absolutely integrable delta function defined on the non-standard extension* $I^*$ *of an interval* $I$ *which is concentrated at a point* $c \in I$. *Then for every bounded continuous real function* $f$ *defined on* $I$ *we have* $\int_{I^*} f^*(x)\delta(x)dx = {}_1 f(c)$.

PROOF. Since $f$ is bounded, say $|f(x)| \leq M$ for all $x \in I$, we have that $\int_{I^*-J} |\delta(x)||f^*(x) - f(c)|dx \leq 2M \int_{I^*-J} |\delta(x)|dx = {}_1 0$, where $J$ is the infinitesimal interval with the property indicated in (iii) of Definition 2.1.

Now by the mean-value theorem of the integral calculus we have that $\int_J \delta(x)$ $(f^*(x) - f(c))dx = (f^*(c+h) - f(c)) \int_J \delta(x)dx$, where h is infinitesimal. Since $\int_J |\delta(x)|dx$ is finite and f is continuous at c, we obtain $\int_J \delta(x)(f(x) - f(c)) = {}_1 0$. This completes the proof of the theorem.

3. Definition of Distributions of Finite Order.

In this section we shall discuss very briefly a realization of generalized functions or distributions as quasi-standard functions. We shall base the discussion on the theory of distributions which was recently given in J. Mikusinski and R. Sikorski, The elementary theory of distributions (I), Rozprawy Matematyczne XII, Warsaw (1957). Following Mikusinski and Sikorski, we shall restrict ourselves to distributions of finite order and of one variable.

In order to facilitate the discussion we shall briefly recall the elements of the theory of distributions of Mikusinski and Sikorski.

Let I be an open but fixed interval with endpoints $a, b (-\infty \leqq a < b \leqq +\infty)$. The set of all real continuous functions defined on I will be denoted by C(I).

Distributions on I, as defined by Mikusinski and Sikorski, are generalized functions in a sense very much similar to Cantor's definition of a real number as a generalized rational number. Thus, they introduce the following definition.

DEFINITION 3.1 (Fundamental sequence of continuous functions). A sequence $\{f_n : n \in N\}$ of real continuous functions defined on I is said to be fundamental if there exist a sequence of real continuous functions $\{\psi_n : n \in N\}$

on I and an integer $k \geq 0$ such that (i) $\psi_n^{(k)}(x) = f_n(x)$ for all $x \in I$, where $\psi_n^{(k)}$ denotes the $k^{th}$ derivative of $\psi_n$ and (ii) the sequence $\{\psi_n : n \in N\}$ converges almost uniformly on I, i.e., the sequence $\{\psi_n : n \in N\}$ converges uniformly on every bounded and closed subinterval of I.

EXAMPLES 1. If $f \in C(I)$, then the sequence $\{f_n : n \in N\}$, where $f_n = f$ for all $n \in N$ is fundamental.

2. Every sequence of continuous real functions on I which converges almost uniformly on I is fundamental. Indeed, in this case we may take in (i) of Definition 3.1 for the sequence $\{\psi_n : n \in N\}$ the sequence $\{f_n : n \in N\}$ and $k = 0$.

3. If the sequence $\{f_n : n \in N\}$ of elements of $C(I)$ is uniformly bounded on I, i.e., $|f_n(x)| \leq M$ for all $x \in I$ and all $n \in N$ and for some constant $M > 0$, and if for some $x_o \in I$ the sequence converges almost uniformly on the open intervals $a < x < x_o$, $x_o < x < b$, then the sequence is fundamental. Indeed, from the hypotheses it follows immediately that we may take the sequence $\{\int_{x_o}^{x} f_n(t)dt : n \in N\}$ as the sequence $\{\psi_n : n \in N\}$ and $k = 1$.

4. From the preceding example it follows now immediately that the sequences $\{1/1+e^{-nx}) : n \in N\}$ , $\{\sqrt{n/2\pi}\, e^{-nx^2/2} : n \in N\}$ are fundamental on R.

5. A sequence of polynomials $\{p_n : n \in N\}$ of degree $< m$ is fundamental on I if and only if it converges uniformly on I. The condition is in an obvious way sufficient (see Example 2). In order to prove that the condition is necessary observe that if a sequence of polynomials $\{p_n : n \in N\}$ of degree $< m$ is fundamental, then there exist a sequence of polynomials $\{q_n : n \in N\}$ on I and an integer $k \geq 0$ such that all polynomials $q_n (n \in N)$ are of degree $< m+k$ and the sequence $\{q_n : n \in N\}$ converges almost uniformly on I.

In order to be able to define certain operations, in particular differentiation, for distribution it is necessary to regard not all fundamental sequences to be different. The theory of distributions, in fact, may be considered as a study of an equivalence relation defined between the elements of the set of all fundamental sequences. This equivalence relation will be introduced in the following defintion.

DEFINITION 3.2 (Equal fundamental sequences). Two fundamental sequences $\{f_n : n\in N\}$ and $\{g_n : n\in N\}$ are called equivalent or equal in the sense of the theory of distributions if there exist sequences $\{\psi_n : n\in N\}$, $\{\chi_n : n\in N\}$ and an integer $k \geq 0$ such that (i) $\psi_n^{(k)}(x) = f_n(x)$, $\chi_n^{(k)}(x) = g_n(x)$ for all $n\in N$ and all $x\in I$ and the sequence $\{\psi_n - \chi_n : n\in N\}$ converges almost uniformly to zero on I.

It is of interest to compare this definition with Cantor's definition of equal Cauchy sequences of rational numbers.

It is easy to see that the relation "equal in the sense of the theory of distributions" is an equivalence relation. The classes of equivalent fundamental sequences are called distributions defined on I. The fundamental sequences given in Example 4 determine the same distribution which is called the Dirac delta distribution. Indeed, if $\{f_n : n\in N\}$ denotes these sequences, then the sequence $\{ \int_{-\infty}^{x} (\int_{-\infty}^{t} f_n(u)du)dt : n\in N\}$ converges almost uniformly to the same function $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x \geq 0$. Consequently, the sequences in Example 4 are equivalent in the sense of the theory of distributions.

We have seen in Example 1 that the constant sequence $\{f : n\in N\}$, where $f\in C(I)$, is fundamental. Hence, it determines a distribution. It is important to observe that different elements $f,g\in C(I)$ determine different distributions. Indeed, if they determine the same distribution, then there exist sequences $\{\psi_n : n\in N\}$, $\{\chi_n : n\in N\}$ and an integer $k \geq 0$ such that $\psi_n^{(k)} = f(x)$, $\chi_n^{(k)}(x) = g(x)$ for all $n\in N$ and all $x\in I$ and $\{\psi_n - \chi_n : n\in N\}$ converges to zero on I almost uniformly. Then the functions $p_n = (\psi_1 - \chi_1) - (\psi_n - \chi_n)$ are polynomials of degree $< k$ since $p_n^{(k)} = 0$. Furthermore, the sequence $\{p_n : n\in N\}$ converges to $\psi_1 - \chi_1$ almost uniformly. Hence, $\psi_1 - \chi_1$ is a polynomial of degree $< k$. Thus its $k^{th}$ derivative f-g is equal to zero, i.e., f and g are <u>not</u> different. This is a contradiction and finishes the proof.

Not all distributions are defined by an element of C(I). The Dirac delta distribution cannot be identified with an element of C(I). Therefore, distributions may be considered as a generalization of continuous functions. This is entirely similar to the generalization of rational numbers as real numbers in Cantor's theory. Indeed, in Cantor's theory a rational number r is identified with the class of fundamental sequences which are equivalent to the constant sequence $\{r\}$.

After this preliminary discussion of Mikusinski and Sikorski's definition of the distribution concept we shall now show in what sense distributions can be represented by means of quasi-standard functions. To this end, we shall first introduce the following definition.

DEFINITION 3.3 ($\sigma$-<u>quasi-standard functions</u>). <u>A quasi-standard function</u> f <u>defined on the non-standard extension</u> $I^*$ <u>of</u> I <u>is called a</u> $\sigma$-<u>quasi-standard function if there exist a sequence</u> $\{g_f(\cdot,n) : n\in N\}$ <u>of real functions on</u> I <u>and an element</u> $\omega_f \in N^*-N$ <u>such that</u> $f(x) = g_f^*(x,\omega_f)$ <u>for all</u> $x\in I^*$.

Then Definition 3.1 and Definition 3.3 justify the following definition.

DEFINITION 3.4 (*Fundamental σ-quasi-standard function*) *A σ-quasi standard function* f *defined on the non-standard extension* $I^*$ *of* I *is called fundamental if it is continuous on* $I^*$ *and if there exist a σ-quasi-standard function* $\psi$ *on* $I^*$ *and an integer* $k \geqq 0$ *such that* $\psi^{(k)}(x) = f(x)$ *for all* $x \in I$ *and* $g^*_\psi(x,\omega) = {}_1g^*_\psi(x,\omega')$ *for all* $x \in I^* \cap M_0$ *and all* $\omega, \omega' \in N^* - N$, *where* $\{g_\psi(\cdot,n) : n \in N\}$ *defines* $\psi$.

The condition $g^*_\psi(x,\omega) = {}_1g^*_\psi(x,\omega')$ for all $x \in I^* \cap M_0$ and all $\omega, \omega' \in N^* - N$ is easily seen to be equivalent to the statement that the sequence $\{g_\psi(\cdot,n) : n \in N\}$ converges almost uniformly on I.

We wish to consider the fundamental σ-quasi-standard functions on $I^*$ as distributions. In order to be able to do this we may not consider every two different fundamental σ-quasi-standard functions to be different in the sense of the theory of distributions. Thus Definition 3.2 justifies the following definition.

DEFINITION 3.4 (*Equivalent fundamental σ-quasi-standard functions*). *Two fundamental σ-quasi-standard functions* $f_1$, $f_2$ *on* $I^*$ *are called equivalent if there exist σ-quasi-standard functions* $\psi$, $\chi$ *on* $I^*$ *and an integer* $k \geqq 0$ *such that* $\psi^{(k)}(x) = f_1(x)$, $\chi^{(k)}(x) =_1 f_2(x)$ *for all* $x \in I$ *and* $g^*_\psi(x,\omega) = {}_1g^*_\psi(x,\omega') = g^*_\chi(x,\omega'') = {}_1g^*_\chi(x,\omega''')$ *for all* $x \in I^* M_0$ *and all* $\omega, \omega', \omega'', \omega''' \in N^* - N$.

*The classes of equivalent fundamental σ-quasi-standard functions on* I *are called distributions.*

This definition is entirely in accordance with Mikusinski and Sikorski's definition of the distribution concept.

Instead of working formally with the classes of equivalent fundamental σ-quasi-standard functions we shall use the fundamental σ-quasi-standard functions and define all the operations on them modulo the equivalence relation. In this connection it is important to remark, that if $f = g_f^*(\cdot,\omega_f)$ is a fundamental $\sigma$-quasi-standard function, then for every $\omega \in N^*-N, g_f^*(\cdot,\omega)$ is a fundamental $\sigma$-quasi-standard function which is equivalent to f. Indeed, this follows immediately from the fact that if $\{f_n : n \in N\}$ is a fundamental sequence, then every subsequence is fundamental and equivalent in the sense of the theory of distributions to $\{f_n : n \in N\}$. Hence, there is no loss of generality if we assume that $\omega_f$ is independent from f. To be more precise, if $\omega_o \in N^*-N$ but fixed, then the set of all fundamental $\sigma$-quasi-standard functions $f = g_f^*(\cdot,\omega_o)$ is a set of representatives of the set of all fundamental $\sigma$-quasi-standard functions in the following sense, that every fundamental $\sigma$-quasi-standard function is equivalent to an element $f = g_f^*(\cdot,\omega_o)$. Thus, we introduce the following notation.

NOTATION. The set of all $\sigma$-fundamental quasi-standard functions f of the forms $g_f^*(\cdot,\omega_o)$, where $\omega_o \in N^*-N$ will be denoted by $D_o$.

Observe, that $D_o$ is a ring. Furthermore, the reader should realize that different elements of $D_o$ may still be equivalent in the sense of the theory of distributions. If $f \in C(I)$, then its non-standard extension $f^*$ on $I^*$ is in $D_o$. In order to distinguish those elements from the general elements of $D_o$, which are usually denoted by f, g, $\cdots$, we shall denote them by $f^*$, $g^*$, $\cdots$ .

## 4. Algebraic Operations on Distributions.

In this section we shall introduce the sum and the difference of two distributions and the product of a distribution with a standard number. The definitions of these operations are generalizations of the same operations in C(I).

To this end, we recall that in the preceding section we have seen that $D_o$ is a ring over R. Consider now the set $J_o$ of all elements f of $D_o$ which are equivalent to zero in the sense of the theory of distributions. Then it is easy to see that if $f \in J_0$, $rf \in J_0$ for all $r \in R$. Furthermore, $f, g \in J_0$ implies that $f \pm g \in J_0$. Hence, $J_0$ is a linear subspace of $D_0$. In fact, the quotient space $D_0/J_0$ is isomorphic to the set of all distributions on I of finite order. Since $D_0/J_0$ is also a linear space over R we have obtained in a natural way the definitions of sum, difference, and multiplication with real numbers for distributions.

The linear subspace $J_0$, however, is <u>not</u> an ideal in $D_0$, hence multiplication for distributions cannot be defined pointwise.. Of course two elements of $D_0$ may be multiplied by one another but equivalent elements may have non-equivalent products. It is possible, however, to multiply a distribution with an infinitely many times differentiable function. Indeed, it is not so very difficult to show that if $\psi$ is a real function defined on I and if $\psi$ is infinitely many times differentiable, then for every $f \in J_0$, $\psi^* f \in J_0$. Hence, if $f \in D_0$, representing a distribution, and $\psi$ is infinitely many times differentiable, then the element $\psi^* f$ determines uniquely a distribution for all $f \in D_0$.

## 5. <u>Derivation of Distributions.</u>

In order to be able to define the derivative of a distribution we need the following simple results.

(i) <u>If f is a fundamental $\sigma$-quasi-standard function which is m-times continuously differentiable and equivalent to zero in the sense of the theory of distributions then $f^{(m)}$ has the same properties.</u>

In the theory of distributions, f equivalent to zero means that there exists a $\sigma$-quasi-standard $\psi$ and an integer $k \geq 0$ such that $f = \psi^{(k)}$ on I and $g^*_\psi(x,\omega) = {}_1 0$ for all $x \in I^* \cap M_0$ and all $\omega \in N^* - N$. Hence, $f^{(m)} = \psi^{(k+m)}$ on I and

$g_\psi(x,\omega) =_1 0$ for all $x \in I^* \cap M_o$ and all $\omega \in N^* - N$, i.e., $f^{(m)}$ is equivalent to zero in the sense of the theory of distributions. Incidentally, we have also shown that $f^{(m)}$ is fundamental.

(ii) <u>For every</u> $f \in C(I)$, <u>there exists a quasi-standard polynomial</u> $p$ <u>such that</u> $f^*(x) =_1 g_p^*(x,\omega)$ <u>for all</u> $x \in I^* \cap M_o$ <u>and all</u> $\omega \in N^* - N$, <u>where</u> $\{g_p(\cdot,n) : n \in N\}$ <u>defines</u> $p$.

Indeed, let $a_n (n \in N)$ be a decreasing sequence of standard numbers and $b_n (n \in N)$ be an increasing sequence of real numbers such that $a = \lim_{n\to\infty} a_n$ and $b = \lim_{n\to\infty} b_n$. From the well-known approximation theorem of Weierstrass it follows that there exists a sequence of polynomials $\{g_p(\cdot,n) : n \in N\}$ on $I$ such that $|f(x) - g_p(x,n)| < \frac{1}{n}$ for all $a_n \leq x \leq b_n$. Hence, $f^*(x) =_1 g_p^*(x,\omega)$ for all $x \in I^* \cap M_o$ and all $n \in N$.

(iii) <u>For every distribution</u> $f$ <u>there exists a fundamental</u> $\sigma$-<u>quasi-standard polynomial</u> $p$ <u>such that</u> $f$ <u>is determined by</u> $p$.

Let $f$ be determined by the fundamental $\sigma$-quasi-standard function $g_f^*(\cdot,\omega_f)$. Then there exists an almost uniformly convergent sequence of continuous functions $\{\psi_n : n \in N\}$ and an integer $k \geq 0$ such that $\psi_n^{(k)}(x) = g_f(x,n)$ for all $x \in I$ and $n \in N$. Let $\psi$ be the limit of the sequence $\{\psi_n : n \in N\}$. Then $\psi$ is continuous; and hence, by (ii), there exists a $\sigma$-quasi-standard polynomial $q$ such that $\psi^*(x) =_1 g_q^*(x,\omega)$ for all $x \in I^* \cap M_o$ and all $\omega \in N^* - N$. Let $p = q^{(k)}$, then it is easy to see that $p$ determines $f$.

DEFINITION 5.1 ($m^{th}$ <u>derivative of a distribution</u>). <u>Let</u> $f$ <u>be a distribution on</u> $I^*$. <u>Then the</u> $m^{th}$ <u>derivative of</u> $f$ <u>is the distribution determined by</u> $p^{(m)}$, <u>where</u> $p$ <u>is a fundamental</u> $\sigma$-<u>quasi-standard function on</u> $I^*$ <u>which determines</u> $f$.

This definition is justified since $p^{(m)}$ is fundamental by (i) and the distribution determined by $p^{(m)}$ does not depend on the representation of f by p.

If a distribution is determined by a m-times continuously differentiable fundamental $\sigma$-quasi-standard function f, then $f^{(m)}$ determines the $m^{th}$ derivative of the distribution.

For continuously differentiable functions the distributional derivative and the derivative in the ordinary sense coincide.

THEOREM 5.1 _If_ f _is a distribution on_ $I^*$, _then_ $f^{(m)} = 0$ _if and only if_ f _is a fundamental_ $\sigma$-_quasi-standard polynomial of degree_ $< m$.

PROOF. Let f be determined by the m-times continuously differentiable fundamental sequence $\{f(\cdot,n) : n \in N\}$. Then $\{f^{(m)}(\cdot,n) : n \in N\}$ is equivalent to zero. Hence, there exist sequence $\{g_\psi(\cdot,n) : n \in N\}$, $\{g_\chi(\cdot,n) : n \in N\}$ and an integer $k \geqq m$ such that $f^{(m)}(x,n) = g_\psi{}^{(k)}(x,n)\ g_\chi{}^{(k)}(x,n) = 0$ for all $x \in I$ and $n \in N$ and $g_\psi^*(x,\omega) = g_\chi^*(x,\omega')$ for all $x \in I^* \cap M_o$ and all $\omega, \omega' \in N^* - N$. Then f - $\psi^{(k-m)}$ is a fundamental $\sigma$-quasi-standard polynomial of degree $< m$. Since $\chi$ is a fundamental $\sigma$-quasi-standard polynomial of degree $< k$ we obtain that $\psi^{(k-m)}$ is such a polynomial of degree $< m$. Hence, f is a fundamental $\sigma$-quasi-standard polynomial degree m; and the proof is finished.

From Theorem 5.1 it follows immediately that $f' = 0$ _implies that_ f _is a constant._

We conclude this section with the following fundamental result in the theory of distributions.

THEOREM 5.2. _Every distribution is the derivative of some order, of a continuous function._

PROOF. Let the distribution f be represented by $g_f^*( \ ,\omega_f)$ on $I^*$. Then there exists a $\psi$ such that $f = \psi^{(k)}$ on I for some $k \geq 0$ and $g_\psi^*(x,\omega) = {}_1g_\psi^*(x,\omega')$ for all $x \in I^* \cap M_0$ and all $\omega, \omega' \in N^* - N$. Let $F(x) = st(g_\psi^*(x,\omega))$ for all $x \in I$. Then f is determined by $(F^*)^{(k)}$.

We shall not continue any further the theory of distributions since that would be beyond the scope of these notes. We merely wanted to point out that the theory of distributions can be realized as the theory of an equivalence relation defined between the elements of the linear space $D_0$ of certain fundamental $\sigma$-quasi-standard functions.

The theory of distributions of infinite order can be treated in exactly the same way. This follows immediately from the definition of distributions of infinite order as given by Mikusinski and Sikorski in section 21 of their paper quoted in section 3 of this chapter.

# CHAPTER 6

# ULTRAPOWERS IN ANALYSIS

## 1. Introduction.

In Chapter 1 we have given a construction of a non-standard model $R^*$ of the system of real numbers R or more precisely of the system of axioms of the theory of totally-ordered fields. From the properties of $R^*$ it follows immediately that this system of axioms is incomplete in the sense of Gödel, i.e., there exists a property $\pi$ such that the system of axioms remains consistent both after adding $\pi$ to it and after adding the property not $(\pi)$. Indeed, R has the property of being archimedean and in $R^*$ the negation of this property holds. Thus the models R and $R^*$ are two models of the axiom system of totally-ordered fields which are not isomorphic.

In general there exist many mutually non-isomorphic models of the arithmetic of the real number system. As we indicated in Chapter 1 different non-standard models of R may be used to analyze more deeply certain parts of analysis. The object of this chapter is to illustrate this method by means of a number of examples. We shall begin in the next section with a presentation of the general theory of the construction of non-standard models of R in the form of ultrapowers.

Some of the contents of this chapter were presented in W. A. J. Luxemburg, Two applications of the method of construction by ultrapowers to analysis, Bull. Amer. Math. Soc., 68 (1962).

## 2. General Non-Standard Models of R in the Form of Ultrapowers.

Let R again denote the set of real numbers. In order to construct a general ultrapower of R we consider the set $R^D$ of all mappings of an infinite set D into R. If $\mathcal{U}$ is an ultrafilter on D, then we introduce in the same way as in Chapter 1 the following relation between the elements of $R^D$. For any elements $A, B \in R^D$ we write $A \equiv_{\mathcal{U}} B$ (read A, B are equal module $\mathcal{U}$) if and only if $\{u : A(u) = B(u) \text{ and } u \in D\} \in \mathcal{U}$. In the same way as in Chapter 1 we can show that $\equiv_{\mathcal{U}}$ is an equivalence relation on the set $R^D$. For each $A \in R^D$, let $a = \{A' : A' \in R^D \text{ and } A \equiv_{\mathcal{U}} A'\}$, the equivalence class of A with respect to the relation $\equiv_{\mathcal{U}}$. The set of the classes of equivalent elements will be denoted by $R^*(D, \mathcal{U})$ or shortly $R^*$ if no confusion can arise. Then the algebraic operations of addition, and multiplication in R can be carried over to $R^*(D, \mathcal{U})$ by means of Definition 4.4 of Chapter 1. In the same way as in Definition 4.5 of Chapter 1, $R^*$ can be ordered. With the introduction of the algebraic operations and order in $R^*$ we can show then as in Chapter 1 that $R^*$ is a totally-ordered field which contains an isomorphic copy of R in the elements of the equivalence classes of the constant mappings of D into R.

In the case that D = N and $\mathcal{U}$ is a <u>free</u> ultrafilter on N we were able to show that $R^*$ is not isomorphic to R or more precisely that $R^*$ is not archimedean. For a general infinite set D one cannot state immediately that $R^*(D, \mathcal{U})$ is a non-standard model of R even if $\mathcal{U}$ is a <u>free</u> ultrafilter. In the following theorem we list two conditions each equivalent to the statement that $R^*(D, \mathcal{U})$ is not isomorphic to R.

THEOREM 2.1. <u>Let</u> D <u>be an arbitrary infinite set and let</u> $\mathcal{U}$ <u>be an ultrafilter on</u> D. <u>Then the following conditions are mutually equivalent</u>.

(a) $R^*(D, \mathcal{U})$ <u>is a proper extension of</u> R.

(b) <u>There exists a mapping</u> $\Omega$ <u>of</u> D <u>into</u> R <u>which is unbounded on every element of</u> $\mathcal{U}$.

(c) $\mathcal{U}$ <u>is a free ultrafilter and there exists a countable family</u> $\{D_n : n\in N\}$ <u>of disjoint non-empty subsets of</u> D <u>such that</u> $\bigcup\{D_n : n\in N\} = D$ <u>and</u> $D_n \notin \mathcal{U}$ <u>for all</u> $n\in N$.

PROOF. (a) implies (b). If $R^* = R^*(D,\mathcal{U})$ is a proper extension of R, then $R^*$ is non-archimedean (Theorem 3.4 of Chap. 1). Hence, there exists an element $a\in R^*$ such that $|a|\geq n$ for all $n\in N$. We conclude that every $A\in a$ is unbounded on every element of $\mathcal{U}$, i.e., (b) holds.

(b) implies (c). If $\mathcal{U}$ is fixed, then there exists an element $u_o\in D$ such that $\{u_o\}\in\mathcal{U}$. But in that case every element of $R^D$ is bounded on the element $\{u_o\}$ of $\mathcal{U}$. This contradicts (b) and proves the first part of (c). In order to prove the second part of (c), let $\Omega$ be a mapping of D into R which is unbounded on every set of $\mathcal{U}$. Then the sets $D_n = \{u : u\in D \text{ and } k < |\Omega(u)| \leq k+1\}$, where k is the first natural number $\geq n$ for which $D_n \neq \emptyset$, $n\in N$, is a non-empty family of non-empty sets such that $D_n \notin \mathcal{U}$ for all $n\in N$ and $D = \bigcup\{D_n : n\in N\}$. Hence, (c) holds.

(c) implies (a). If (c) holds, let $\Omega$ be the following mapping of D into R. For all $n\in N$, $u\in D_n$ implies that $\Omega(u) = n$. Then if $\omega$ is that element of $R^*$ which is determined by $\Omega$ we have that $\omega \geq n$ for all $n\in N$, i.e., $R^*$ is non-archimedean or in other words $R^*$ is a proper extension of R. This proves (a) and completes the proof of the theorem.

REMARK. In connection with the preceding theorem the following question arises: <u>Given an infinite set</u> D. <u>Does there exist a free ultrafilter</u> $\mathcal{U}$ <u>on</u> D <u>such that for every countable family</u> $\{D_n : n\in N\}$ <u>of disjoint non-empty subsets of</u> D, $\bigcup\{D_n : n\in N\} = D$ <u>implies that for some index</u> n (and hence for only one), $D_n \in \mathcal{U}$? In the case that D = N we have shown that the answer is no. For general infinite sets this question is equivalent to Ulam's two-valued measure problem. By the latter we mean the following problem: Given an infinite set D. Does there exist a $\{0,1\}$-valued measure on D which is countably additive and which vanishes on every one point set and which takes on the value 1 on D? or equivalently: Does there exist a mapping $\mu$ of the power set p(D) of D into the set $\{0,1\}$ such that (i) for all $u \in D$, $\mu(\{u\}) = 0$; (ii) $\mu(D) = 1$ and (iii) for every countable family $\{D_n : n \in N\}$ of mutually disjoint subsets of D we have $\mu(\cup D_n) = \Sigma \mu(D_n)$.

In order to prove that these two questions are equivalent assume first that $\mu$ is a countably additive measure on p(D) with properties (i) and (ii). Then we introduce the following collection $\mathcal{U}$ of subsets of D. We have $X \in \mathcal{U}$ if and only if $\mu(X) = 1$. Then it is easy to see that $\mathcal{U}$ is a free ultrafilter on D with the required property. The converse is proved similarly.

It is evident that this is a question about cardinals. We call a cardinal $\alpha$ measurable if there exists a set D of cardinal $\alpha$ which admits a $\{0,1\}$-valued measure having the properties (i), (ii) and (iii) stated above. In the other case, we call a cardinal non-measurable. Our problem then is to find all non-measurable cardinals. We have shown incidentally that card(N) is non-measurable. It is not known whether every infinite cardinal is non-measurable. Recent research on this problem indicates, however, that measurable cardinals if they exist have to be tremendously large. For an up-to-date account of the measure problem we refer the reader to the paper: A. Tarski, Some problems and results relevant to the foundations of set theory, Proc. Int. Congr. Logic, Methodology and Philosophy of Science, Stanford (1960).

Finally, we would like to point out that, in contrast to the above problem, on every infinite set there exist free ultrafilters with the property given in (c) of Theorem 2.1. Indeed, if D is an infinite set and $\{D_n : n \in N\}$ is a partition of D in non-empty sets, then consider the filter F on D which is generated by the complements of the <u>finite</u> unions of sets of the given family $\{D_n : n \in N\}$. Then any ultrafilter $\mathcal{U}$ on D containing F has the property that $D_n \notin \mathcal{U}$ for all $n \in N$.

Assume now that D is an infinite set and $\mathcal{U}$ is a free ultrafilter on D such that $R^*(D, \mathcal{U})$ is a proper extension of R. In this case we introduce as in section 5 of Chapter 1 the notions of infinitesimal, finite and infinitely large number. Again we shall denote the set of all infinitesimals and all finite elements of $R^* = R^*(D, \mathcal{U})$ by $M_1 = M_1(D, \mathcal{U})$ and $M_o = M_o(D, \mathcal{U})$ respectively. Then $M_o$ is an integral domain and $M_1$ is a maximal ideal in $M_o$. Furthermore, the field $M_o/M_1$ is isomorphic to R. Those results can be easily verified by the reader. The homomorphism of $M_o$ onto R with kernel $M_1$ will again be denoted by "st" and again called the "standard part" homomorphism. We shall again call the elements of $R^*$ which are not in R the non-standard elements of $R^*$ and we shall often refer to the elements of R as the standard elements of $R^*$, in this case also. Furthermore, we shall write again $a =_1 b$ if a and b are infinitely close to each other. With this terminology we have again that if $a \in M_o$, st(a) is the unique standard number infinitely close to a.

REMARK. Incidentally, whether $R^*$ is a proper extension of R or not, the sets $M_o$ and $M_1$ can be introduced in the same way. If $R^*$ and R are isomorphic, however, then $M_o = R^*$ and $M_1 = \{o\}$ and the "standard part homomorphism" is the identity.

Our next task will be to show that if $R^*(D, \mathcal{U}) \neq R$, the theory developed in the preceding chapters holds in $R^*(D, \mathcal{U})$ as well. More precisely, we mean that the formulation and characterization of the basic notions and principles of analysis in non-standard language are independent from the ultrapowers $R^*(D,\mathcal{U})$ of R we use, as long as they are proper extensions of R. We shall see that this fact depends essentially on Theorem 2.1. We shall illustrate this first by means of an example. It is evident what we mean in the general case by the non-standard extension of a subset of R and the non-standard extension of a binary relation in R. Now we showed in Chapter 1, that a subset S of R is infinite if and only if $S \neq S^*$. We shall give now a proof of this statement for the general case. Assume first that S is finite, say $S = (s_1, \ldots, s_n)$. If $a \in S^*$, then $\{u : u \in D \text{ and } A(u) \in S\} \in \mathcal{U}$, where $A \in a$. Since $\mathcal{U}$ is an ultrafilter, there exists an index i, $1 \leq i \leq n$, such that $\{u : u \in D \text{ and } A(u) = s_i\} \in \mathcal{U}$. Hence, a is a standard element, i.e., $a \in S$. Conversely, assume that S is infinite. Then there exists an injection $\psi$ of N into S. Now we define the following mapping A of D into S. If $u \in D_n$, then $A(u) = \psi(n)$ for all $n \in N$, where $\{D_n : n \in N\}$ is a partition of D into non-empty disjoint sets such that $D_n \notin \mathcal{U}$ for all $n \in N$. Then, if $a \in R^*$ is such that $A \in a$, we have that $a \in S^*$ but $a \notin S$.

This example illustrates precisely how to use property ($\dot{a}$) of Theorem 2.1 in the case that $R^*(D,\mathcal{U}) \neq R$.

In general, the fact that the formulation and characterization of the basic notions and principles of analysis in non-standard language are independent from the ultrapowers $R^*(D,\mathcal{U})$ of R we use, as long as they are proper extensions of R, follows easily from the following theorem.

THEOREM 2.2. Let D be an infinite set and let $\mathcal{U}$ be an ultrafilter on D such that the ultrapower $R^*(D,\mathcal{U})$ is a proper extension of R. Then there exists a free ultrafilter $\mathcal{U}'$ on N such that $R^*(N,\mathcal{U}')$ is isomorphic to a subfield of $R^*(D,\mathcal{U})$.

PROOF. From the hypothesis that $R^*(D,\mathcal{U})$ is a proper extension of R it follows that there exists a partition $\{D_n : n\epsilon N\}$ of D consisting of non-empty subsets of D such that $D_n \notin \mathcal{U}$ for all $n\epsilon N$ (Theorem 2.1). Let E be a subset of D. Then by $E_s$ we denote the set $\cup\{D_n : D_n \cap E \neq \emptyset\}$, i.e., $E_s$ is the saturation of E with respect to the partition $\{D_n : n\epsilon N\}$ of D or the equivalence relation defined by the partition $\{D_n : n\epsilon N\}$. Every saturated subset $E_s$ of D defines in a unique way a subset of N in the form of the set of all indices n such that $D_n \subseteq E_s$. The sets so defined on N by the saturated elements of $\mathcal{U}$ form a free ultrafilter $\mathcal{U}'$ on N. The fact that $R^*(N,\mathcal{U}')$ is isomorphic to a subfield of $R^*(D,\mathcal{U})$ is then easily verified by considering the set of those elements of $R^D$ which are constant on the sets $D_n (n\epsilon N)$. This completes the proof of the theorem.

In summary, the above theorem implies that the non-standard interpretation of analysis given in the preceding chapters is the restriction of a similar interpretation in $R^*(D,\mathcal{U})$ to $R^*(N,\mathcal{U}')$.

We conclude this section with a simple theorem which is frequently used in the remainder of this chapter.

Let X be any set. A non-empty family $\{X_a : a\epsilon A\}$ of subsets of X is said to have the finite intersection property if for every finite subset F of A we have $\cap\{X_a : a\epsilon F\} \neq \emptyset$. Hence, in particular, $X_a \neq \emptyset$ for all $a\epsilon A$.

THEOREM 2.3 *Let X be any set and let $\{X_a : a\epsilon A\}$ be a non-empty family of subsets of X with the finite intersection property. Then the family $\{X_a : a\epsilon A\}$ is included in some ultrafilter $\mathcal{U}$ on X.*

PROOF. Let $\mathcal{F}$ be a set of subsets of X which is defined as follows: $Y\epsilon\mathcal{F}$ implies there exists a finite subset F(Y) of A such that $Y \supseteq \cap\{X_a : a\epsilon F(Y)\}$.

Then $\mathcal{F}$ is a filter on X. Indeed, since the family $\{X_a : a \in A\}$ has the finite intersection property it follows that $\emptyset \notin \mathcal{F}$. Furthermore, $Y \in \mathcal{F}$ and $Y' \supseteq Y$ obviously implies that $Y' \in \mathcal{F}$. Finally, if $Y, Y' \in \mathcal{F}$, then $Y \cap Y' \supseteq \cap \{X_a : a \in F(Y) \cup F'(Y)\}$ implies that $Y \cap Y' \in \mathcal{F}$. Hence, $\mathcal{F}$ satisfies the filter properties $(F_1)$, $(F_2)$ and $(F_3)$ (Definition 2.1, Chapter 1), i.e., $\mathcal{F}$ is a filter on X. We conclude from Theorem 2.1 of Chapter 1 that there exists an ultrafilter $\mathcal{U}$ on X which includes $\mathcal{F}$. This completes the proof.

3. The Hahn-Banach Extension Theorem.

As a first application of the use of general non-standard models of R in analysis we shall give a new proof of the very important Hahn-Banach extension theorem.

THEOREM 3.1 (Hahn-Banach). *Let* E *be a real linear space and let* p *be a sublinear functional defined on* E, *i.e., a mapping* p *of* E *into* R *such that* $p(x+y) \leq p(x) + p(y)$ *for all* $x, y \in E$ *and* $p(tx) = tp(x)$ *for all* $x \in E$ *and all* $t \geq 0$. *If* f *is a real linear functional defined on a linear subspace* G *of* E *such that* $f(x) \leq p(x)$ *for all* $x \in G$, *then there exists a real linear functional* F *on* E *such that* $F(x) = f(x)$ *for all* $x \in G$ *and* $F(x) \leq p(x)$ *for all* $x \in E$.

PROOF. We note first that if $z \in E$ but $z \notin G$, then there exists a real linear functional g, defined on the linear subspace G' of all elements $x = v + \lambda z$ ($v \in G$ and $\lambda$ real), such that $g(x) = f(x)$ on G and $g(x) \leq p(x)$ on G'. Indeed, observe that the representation $x = v + \lambda z$ for an element $x \in G'$ is unique. The problem is, therefore, to find a suitable value for $\alpha = g(z)$. Since the definition $g(v + \lambda z) = f(v) + \lambda\alpha$ for every $x = v + \lambda z \in G'$ extends f linearly to G', the requirement concerning $\alpha$ is that $g(v + \lambda z) = f(v) + \lambda\alpha \leq p(v + \lambda z)$ for every $v \in G$ and all real $\lambda \neq 0$. For $\lambda > 0$ we set $v = \lambda w$

and we write the inequality in the form $f(w) + \alpha \leqq p(w + z)$, which holds for all $w \in G'$. For $\lambda < 0$ we set again $v = \lambda w$, and now dividing by $-\lambda$, the inequality becomes $-f(w) - \alpha \leqq p(-w-z)$ for all $w \in G'$. Hence, the requirement is that

$$-p(-w_1-z) - f(w_1) \leqq \alpha \leqq p(w_2+z) - f(w_2)$$

for all $w_1, w_2 \in G$. But $f(w_2)-f(w_1) = f(w_2-w_1) \leqq p(w_2-w_1) = p\{(w_2+z) + (-w_1-z)\}$ $\leqq p(w_2+z) + p(-w_1-z)$, so $- p(-w_1-z) - f(w_1) \leqq p(w_2+z) - f(w_2)$ for all $w_1, w_2 \in G$. Thus $P = \sup_{w_1 \in G}\{- p(-w_1-z) - f(w_1)\} \leqq \inf_{w_2 \in G}\{p(w_2+z) - f(w_2)\} = Q$. It follows that $\alpha$ can be taken as any number satisfying $P \leqq \alpha \leqq Q$.

Having established this result we proceed as follows: Let $\{f_u : u \in D\}$ be the family of all linear functionals which are defined on some linear subspace of E which contains G and which have the following properties: $f_u(x) = f(x)$ for all $x \in G$ and $f_u(x) \leqq p(x)$ for all $x$ in the domain of $f_u$. It is evident that $D \neq \emptyset$. For every $x \in E$ we denote by $D_x$ the set of all $u \in D$ such that $x \in$ domain $(f_u)$. From the result above it follows that $D_x \neq \emptyset$ for all $x \in E$. Furthermore, the family $\{D_x : x \in E\}$ of non-empty subsets of D has the finite intersection property, i.e., if $x_1, \cdots, x_n$ are elements of E, then $\bigcap_{i=1}^{n} D_{x_i} \neq \emptyset$. Indeed, apply the above result successively to the elements $x_1, \cdots, x_n$. Now, let $\mathcal{U}$ be an ultrafilter on D which contains the family $\{D_x : x \in E\}$ (Theorem 2.3). Then we define the following mapping $\tilde{f}$ of E into $R^*(D, \mathcal{U})$. If $x \in E$, then $\tilde{f}(x)$ is that element of $R^*(D, \mathcal{U})$ which is determined by an element $A \in R^D$ such that $A(u) = f_u(x)$ for all $u \in D_x$. Then it is easy to see that $\tilde{f}$ is a linear transformation of E into $R^*(D, \mathcal{U})$ (consider $R^*$ as a linear space over R) and that $\tilde{f}$ has the following two properties: (i) $\tilde{f}(x) = f(x)$ for all $x \in G$ and (ii) $\tilde{f}(x) \leqq p(x)$ for all $x \in E$. From (ii) it follows that $- p(-x) \leqq \tilde{f}(x) \leqq p(x)$ for all $x \in E$, i.e., $\tilde{f}(x)$ is finite for all

$x \in E$. Hence, $F = st(\tilde{f})$ is the required linear functional. This completes the proof of the theorem.

REMARK  It may be of interest to point out that the use of non-standard arguments in the above proof of the Hahn-Banach extension theorem eliminates the use of Zorn's lemma. In fact, we derive the Hahn-Banach extension theorem directly from the apparently weaker hypothesis that every proper filter is contained in an ultrafilter, or equivalently, the prime ideal theorem for Boolean algebras. Whether the converse holds, i.e., whether the Hahn-Banach extension theorem implies the prime ideal theorem for Boolean algebras, seems to be an open problem.

## 4. Tarski's Extension Theorem for Measures on Boolean Algebras.

Let B be a Boolean algebra. A mapping m of B into R is called a measure on B if m has the following properties:

($m_1$) $m(o) = 0$ ; and $o \leq m(a) \leq +\infty$ for all $a \in B$.

($m_2$) If $a, b \in B$ and $a \leq b$, then $m(a) \leq m(b)$ (m is monotone).

($m_3$) If $a, b \in B$ and $a \wedge b = 0$, then $m(a \vee b) = m(a) + m(b)$ (m is additive).

It is evident that $m(o) = 0$ and ($m_2$) imply that $m(a) \geq 0$ for all $a \in B$. Furthermore, we call a measure m a finite measure if $m(1) < +\infty$.

From ($m_3$) it follows immediately that if $a_1, \cdots, a_n$ is a finite family of elements of the Boolean algebra which are mutually disjoint ($a_i \wedge a_j = 0$ if $i \neq j$), then $m(a_1 \vee a_2 \vee \cdots \vee a_n) = \sum_{i=1}^{n} m(a_i)$.

In A. Tarski, Une contribution a la théorie de la mesure, Fund. Math., 15, 42-50 (1930), we find the following interesting theorem, which we shall prove here by means of non-standard arguments.

THEOREM 4.1 (Tarski). Every finite measure $m_o$ defined on a subalgebra $B_o$ of a Boolean algebra $B$ can be extended to a measure $m$ on $B$ such that the set of values of $m$ is contained in the closure of the set of values of $m_o$.

PROOF. Let $B_o$ be a proper subalgebra of a Boolean algebra $B$ and let $m_o$ be a finite measure defined on $B_o$. We shall denote by $S_o$ the set of all values of $m_o$. Then, by hypothesis, $S_o$ is a bounded subset of $R$.

We shall first show, analogously to the proof of the Hahn-Banach extension theorem, the following essential result.

If $a_o \notin B_o$ is an element of $B$, then $m_o$ can be extended to a finite measure $m_1$ on the subalgebra $B_1$ generated by $B_o$ and the element $a_o$ such that the range of $m_1$ is contained in the closure of the range of $m_o$.

To this end, recall that $B_1$ is the set of all elements $a \in B$ which can be expressed in the form $a = (a_1 \wedge a_o) \vee (a_2 \wedge a_o')$, where $a_1, a_2 \in B$ and $a_o'$ is the complement of $a_o$ in $B$ (R. Sikorski, Boolean Algebras, Ergebnisse 25, section 4(3)).

We then define

$m_1(a_1 \wedge a_o) = \inf(m(u) : u \geq a_1 \wedge a_o \text{ and } u \in B_o)$

$m_1(a_2 \wedge a_o') = \sup(m(v) : v \leq a_2 \wedge a_o' \text{ and } v \in B_o)$

and claim that $m_1(a) = m_1(a_1 \wedge a_o) + m_1(a_2 \wedge a_o')$ for some representation of an element $a \in B_1$ is the required extension. We shall now briefly indicate the necessary steps for proving this statement. First of all, if $b \in B_o$, observe that for any $u \in B_o$ such that $u \geq a_o \wedge b$ we have $u' \wedge b \leq a_o' \wedge b$. Hence,

$m_1(b) = m_o(b)$ for all $b \in B_o$. In order to show that $m_1$ is additive (the other two properties are evident) it is enough to show that for all $a_1, a_2 \in B_o$, $(a_1 \wedge a_o) \wedge (a_2 \wedge a_o) = 0$ implies that $m_1\{(a_1 \wedge a_o) \vee (a_2 \wedge a_o)\} = m_1(a_o \wedge a_1) + m_1 (a_o \wedge a_2)$. It is easy to see that $m_1(a_o \wedge a_1) + m_1 (a_o \wedge a_2) \geq m_1 \{(a_1 \wedge a_o) \vee (a_2 \wedge a_o)\}$. To prove the converse inequality, we observe that for any $w \in B_o$ such that $w \geq (a_o \wedge a_1) \vee (a_o \wedge a_2) = a_o \wedge (a_1 \vee a_2)$ we have $m(w) \geq m(w \wedge a_1 \wedge a_2') + m(w \wedge a_2)$. From $w \wedge a_1 \wedge a_2' \geq a_o \wedge a_1$ it follows that $m(w \wedge a_1 \wedge a_2') \geq m_1(a_o \wedge a_1)$; in the same way we have $m(w \wedge a_2) \geq m_1(a_o \wedge a_2)$. Thus $m_1\{(a_o \wedge a_1) \vee (a_o \wedge a_2)\} \geqq m_1(a_o \wedge a_1) + m_1(a_o \wedge a_2)$, which shows that $m_1$ is a measure on $B_1$ which extends $m_o$. That the range of $m_1$ is contained in the closure of the range of $m_o$ follows immediately from the definition of $m_1$.

In order to complete the proof of Tarski's theorem we proceed as follows: Let $\{m_u : u \in D\}$ be the family of all measures defined on subalgebras of B which extend $m_o$ and which have their range in the closure $\overline{S}_o$ of $S_o$. It is obvious that $D \neq \emptyset$. For every $a \in B$, let $D_a$ be the set of all $u \in D$ such that $a \in$ domain $(m_u)$. Then $D_a \neq \emptyset$ for all $a \in B$ and, in fact, the family $\{D_a : a \in B\}$ of non-empty subsets of B has the finite intersection property (apply the above result a finite number of times). Let $\mathcal{U}$ be an ultrafilter on D which contains the family $\{D_a : a \in B\}$ (Theorem 2.3). Then we define the following mapping $\tilde{m}$ of B into $R^*(D, \mathcal{U})$. If $a \in B$, then $\tilde{m}(a)$ is that element of $R^*(D, \mathcal{U})$ which is determined by an element $M \in D^R$ such that $M(u) = m_u(a)$ for all $u \in D_a$. Then it is easy to verify that $\tilde{m}$ is a $R^*$-valued measure on B with the properties: $\tilde{m}(a) = m_o(a)$ for all $a \in B_o$ and $\tilde{m}(a) \in (\overline{S}_o)^*$ for all $a \in B$, where $(\overline{S}_o)^*$ is the non-standard extension of $\overline{S}_o$. Since $S_o$ is bounded, we have that $(\overline{S}_o)^* \in M_o$ (Theorem 7.5 of Chap. 1) and $st\{(\overline{S}_o)^*\} = \overline{S}_o$ (Theorem 1.2 of Chap. 3). Hence, $m = st(\tilde{m})$ is the required extension; and the proof is finished.

The above proof of Tarski's theorem shows, as was the case with the Hahn-Banach extension theorem, that it is an immediate consequence of the ultrafilter hypothesis, i.e., every proper filter is contained in an ultrafilter. As we pointed out in the preceding section we do not know whether the Hahn-Banach extension theorem, conversely, implies the ultrafilter hypothesis. In this case we can show very easily, however, that the converse, indeed, holds, i.e., we have the following theorem.

THEOREM 4.2. Tarski's extension theorem for finite measures implies the ultrafilter theorem, i.e., that every (proper) filter is contained in an ultrafilter.

PROOF. Let X be an arbitrary (infinite) set and let $\mathfrak{F}$ be a filter on X. For this occasion, we denote the Boolean algebra of all subsets of X by B. The subalgebra of all elements $E \in B$ such that either $E \in \mathfrak{F}$ or its complement $E' = X - E \in \mathfrak{F}$ (they cannot both belong to $\mathfrak{F}$ since $\mathfrak{F}$ is not degenerated) will be denoted by $B_o$. That $B_o$ is a subalgebra of B is easy to verify. For every $E \in B_o$ we set $m(E) = 1$ if $E \in \mathfrak{F}$ and $m(E) = 0$ if $E \notin \mathfrak{F}$. Then $m_o$ is a measure on $B_o$. Indeed, $(m_1)$ follows from the fact that $X \in \mathfrak{F}$; $(m_2)$ follows immediately from the property $(F_1)$ of $\mathfrak{F}$ (Definition 2.1 of Chapter 1); finally $(m_3)$ can be proved as follows: Assume that $E_1, E_2 \in B_o$ and $E_1 \cap E_2 = \emptyset$. Then either $E_1 \cup E_2 \in \mathfrak{F}$ or $(E_1 \cup E_2)' \in \mathfrak{F}$. If $E_1 \cup E_2 \notin \mathfrak{F}$, then by property $(F_1)$ of $\mathfrak{F}$, $E_1 \notin \mathfrak{F}$ and $E_2 \notin \mathfrak{F}$ and hence $m(E_1 \cup E_2) = m(E_1) + m(E_2)$. If $E_1 \cup E_2 \in \mathfrak{F}$, then $E_1' \in \mathfrak{F}$ and $E_2' \in \mathfrak{F}$ is impossible. Indeed $(E_1 \cup E_2)' = E_1' \cap E_2'$ implies, by property $(F_2)$ of $\mathfrak{F}$, that $(E_1 \cup E_2)' \in \mathfrak{F}$. Hence, we have either $E_1 \in \mathfrak{F}$ or $E_2 \in \mathfrak{F}$. This proves that $m_o$ is a measure on $B_o$. We conclude from Tarski's extension theorem that there exists a finite measure m on B which extends $m_o$ and which has its values in the closure of the range of $m_o$. Since $m_o$ is $\{0,1\}$-valued it follows that m is

$\{0,1\}$-valued also. Now, let $\mathcal{U} = \{E : E \subset X \text{ and } m(E) = 1\}$. Then it follows immediately that $\mathcal{F} \subseteq \mathcal{U}$ and $\mathcal{U}$ is a filter on X. We shall prove that $\mathcal{U}$ is an ultrafilter on X. According to Theorem 2.4 of Chap. 1 we have to show that for every $E \subseteq X$ either $E \in \mathcal{U}$ or $E' \in \mathcal{U}$. Observe that $E \cup E' = X$; and hence, by property $(m_3)$ of m, we have that $m(E) + m(E') = m(X) = 1$, i.e., either $m(E) = 1$ or $m(E') = 1$, or equivalently, $\mathcal{U}$ is an ultrafilter. This completes the proof of the theorem.

REMARKS 1. We do not know how to derive the Hahn-Banach extension theorem directly from Tarski's extension theorem.

2. In the same way as in the above proof of Tarski's extension we can use non-standard arguments to prove the following theorem of J. L. Kelley (see J. L. Kelley, Measures on Boolean Algebras, Pacific Journ. of Math., 9, 1165-1176 (1959)).

THEOREM 4.3 (Kelley). *Let* B *be a Boolean algebra and let* p *be a non-negative monotone real function on* B *such that* $p(a) + p(b) \leq p(a \vee b) + p(a \wedge b)$ *for all* $a, b \in B$, *and let* $m_o$ *be a measure on a subalgebra* $B_o$ *of* B *such that* $m_o(a) \leq p(a)$ *for all* $a \in B_o$. *Then there exists a measure* m *on* B *such that* $m(a) = m_o(a)$ *for all* $a \in B_o$ *and* $m(a) \leq p(a)$ *for all* $a \in B$.

I do not know whether this result is equivalent to the ultrafilter theorem.

## 5. Nikodým's Theorem about the Existence of Strictly Positive Non-Archimedean Measures on Boolean Algebras.

Let B be a Boolean algebra. A measure m on B is called *strictly positive* or *effective* if $m(a) = 0$ if and only if $a = 0$. Every Boolean algebra admits in a *trivial way*, by defining $m(a) = +\infty$ for all $a \in B$ and $a \neq 0$, a strictly positive measure. If, on the other hand, we require the measure to be finite, then the problem of the existence of an effective measure is quite different. In fact, there exist Boolean algebras (even complete non-atomic Boolean algebras) which do not admit any finite effective measure. The following example of such an algebra is due to J. Dixmier.

Let X be an infinite set of larger cardinal than the cardinal of N. Let B be again the Boolean algebra of all subsets of X and let J be the set of all subsets E of X such that $\mathrm{card}(E) \leq \mathrm{card}(N)$. Then J is an ideal in B. Let $B' = B/J$. We shall prove that $B'$ does not admit any effective finite measure. In order to see this, assume that $B'$ does admit such a measure m. Let $a \in B'$ and $a \neq o$, then $m(a) > o$. Furthermore, if the subset A of X is in the class a, then $\alpha = \mathrm{card}\,(A) > \mathrm{card}(N)$. From the fact that $\alpha\alpha = \alpha$ it follows that there exists a family $\{A_p : p \in P\}$ of mutually disjoint subsets of X such that $\mathrm{card}(A_p) > \mathrm{card}(N)$ for all $p \in P$ and $\mathrm{card}(P) = \alpha$. Let $a_p$ be the class of the element $A_p$ $(p \in P)$. Then $a_p \wedge a_q = 0$ for all $p \neq q$. Furthermore, we have that $m(a_p) > 0$ for all $p \in P$. Since $\mathrm{card}(P) = \alpha > \mathrm{card}(N)$ there exists a positive number $\delta > 0$ such that $m(a_p) > \delta$ for more than countably many $p \in P$. Hence, for every $n \in N$ we have that $m(a_{p_i}) > \delta$, $i=1, 2, \cdots, n$, implies $n\delta \leq m(1)$. This contradicts $\delta > 0$ and the proof is completed.

In the light of this result, the following theorem of O. Nikodým, which is contained in the papers: O. Nikodým. On extension of a given finitely additive field-valued non-negative measure, on a finitely additive Boolean tribe, to another tribe more ample, Rend. Sem. Mat. Univ. di Padova 26, 232-327 (1956), O. Nikodým, Sur le mesure non-archimédienne effective sur une tribu de Boole arbitraire, C. R. Acad. Sci. Paris 251, 2113-2115 (1960), is of interest. We shall prove it here in a much simpler way using non-standard arguments.

THEOREM 5.1 (Nikodým). For every Boolean algebra B there exists a totally-ordered field F, which is in general non-archimedean, such that B admits a strictly positive F-valued finite measure.

PROOF. Our proof depends on the following fundamental result of M. H. Stone (see R. Sikorski, Boolean Algebras, Theorem 6.1). If B is a Boolean algebra, then for every $a \in B$ and $a \neq 0$ there exists a $\{0,1\}$-valued Boolean homomorphism h such that $h(a) = 1$. Incidentally, this result is equivalent to the ultrafilter hypothesis. If we observe that a Boolean two-valued homomorphism is a two-valued measure, then Stone's results implies that for every $a \in B$ and $a \neq 0$ there exists a finite measure m such that $m(a) \neq 0$. Furthermore, since the sum of a finite number of measure is again a measure we have, in fact, that for every finite subset $a_1, \cdots, a_n$ of non-zero elements of B there exists a finite measure m on B such that $m(a_i) \neq 0$ for all $i = 1, 2, \cdots, n$. Having established this result, we proceed then as follows: Let $\{m_u : u \in D\}$ be the family of all non-zero real measures on B. There is no loss in generality if we assume that $m_u(1) = 1$ for all $u \in D$. Indeed, if $m(1) \neq 1$, then by considering the measure $m' = m/m(1)$ we obtain a measure $m'$ such that $m'(1) = 1$. For every $a \in B$ and $a \neq 0$, we denote by $D_a$ the set of all $u \in D$ such that $m_u(a) \neq 0$. From the result above, which we deduced from the theorem of Stone about the existence of sufficiently many two-valued homomorphisms, it follows that the family $\{D_a : a \in B \text{ and } a \neq 0\}$ of non-empty subsets of D has the finite intersection property. Let $\mathcal{U}$ be an ultrafilter on D which contains the family $\{D_a : a \in B \text{ and } a \neq 0\}$ (Theorem 2.3). Then $F = R^*(D, \mathcal{U})$ is a totally-ordered field which is in general non-archimedean. Then we define the following mapping $\tilde{m}$ of B into F. If $0 \neq a \in B$, then $\tilde{m}(a)$ is that element of F for which there exists an element $M \in \tilde{m}(a)$ such that $M(u) = m_u(a)$ for all $u \in D_a$; and we define $\tilde{m}(o) = 0$. Then, by construction, $\tilde{m}$ has the following properties: (i) $\tilde{m}(a) = 0$ if and only if $a = 0$, i.e., $\tilde{m}$ is strictly positive and (ii) $\tilde{m}$ is an F-valued measure on B such that $\tilde{m}(1) = 1$. This completes the proof of the theorem.

REMARKS. 1. If B does not admit a strictly positive real finite measure, then the totally-ordered field F constructed in the proof of the preceding theorem is a proper extension of R and hence, $\tilde{m}(a)$ is infinitesimal for at least one element $0 \neq a \in B$.

2. If in place of the family of all normalized measures on B we take the set of all non-zero two-valued measures and their averages we obtain, in addition, that the totally-ordered field F for which Nikodým's theorem holds may be a proper extension of the field of rational numbers Q.

## 6. The Heine-Borel Covering Theorem.

In this section we shall give a non-standard proof of the classical Heine-Borel covering theorem.

THEOREM 6.1. (Heine-Borel) A subset S of $R_p$ $(p \geq 1)$ is bounded and closed if and only if S is compact, i.e., every open covering of S has a finite subcovering.

PROOF. If $S \subset R_p$ is compact, then obviously every ultrafilter on S is convergent. Hence, if $\mathcal{U}$ is an ultrafilter on S and $R_p^* = R_p^S/\mathcal{U}$, then $S^* \subset M_0^p$ and $st(S^*) = S$. This shows that the condition is sufficient (See Theorems 2.1 and 4.1 of Chapter 4).

In order to prove that the condition is necessary we shall assume that $\{O_\lambda : \lambda \in \Lambda\}$ is an open covering of S. If this covering does not contain a finite subcovering of S, then for every finite subset $F \subset \Lambda$ we have that

$A_F = S - \cup(O_\lambda : \lambda \in F)$ is not empty. Furthermore, the non-empty family of non-empty sets $\{A_F : F \subset \Lambda \text{ and } F \text{ is finite}\}$ has the finite intersection property. Hence, by Theorem 2.3, there exists an ultrafilter $\mathcal{U}$ on S which contains the family $\{A_F : F \subset \Lambda \text{ and } F \text{ is finite}\}$. Consider then the ultrapower $R_p^* = R_p^S/\mathcal{U}$. Let E denote the identity mapping of S onto S, i.e., $E(x) = x$ for all $x \in S$, and let $e \in R_p^*$ be such that $E \in e$. Then $e \in S^*$ and $e \notin O_\lambda^*$ for all $\lambda \in \Lambda$. Since S is bounded and closed there exists an element $s \in S$ and an infinitesimal $h \in M_1^p$ such that $e = s + h$. For some $\lambda$, $s \in O_\lambda$. Hence, $O_\lambda$ being open we have that $e = s + h \in O_\lambda^*$. This contradicts the definition of $\mathcal{U}$ and the proof is finished.

REMARKS. 1. It is not without interest to observe that under the hypothesis that the covering $\{O_\lambda : \lambda \in \Lambda\}$ has no finite subcovering it follows that for any ultrafilter $\mathcal{U}$ of subsets of S which contains the elements of the family $\{A_F : F \subset \Lambda \text{ and } F \text{ is finite}\}$ the ultrapower $R_p^* = R_p^S/\mathcal{U}$ of $R_p$ is a proper extension of $R_p$. Indeed, since $R_p$ and hence S have the Lindelöf property it follows that there exists a countably infinite subset $\Lambda_0 \subset \Lambda$ such that the family $\{O_\lambda : \lambda \in \Lambda_0\}$ covers S, then $O_\lambda \notin \mathcal{U}$ for all $\lambda \in \Lambda_0$ implies that $R_p^*$ is a proper extension of $R_p$ (See Theorem 2.1).

2. From the above method of proof we deduce immediately the following abstract covering theorem.

THEOREM 6.2. *Let X be an arbitrary set and let $\{O_\lambda : \lambda \in \Lambda\}$ be a covering of X. Then this covering contains a finite subcovering of X if and only if every free ultrafilter on X contains an element of the covering.*